

Université de Montréal

# **Étude sur la reconnaissance de l'ubiquitine par les domaines de transactivation acides des activateurs de transcription**

par

Mathieu Lussier-Price

Département de biochimie

Faculté de Médecine

Mémoire présenté à la Faculté de médecine

en vue de l'obtention du grade de M.Sc.

en biochimie

option biologie structurale

26 mars, 2014

© Mathieu Lussier-Price, 2014

Université de Montréal

Faculté des études supérieures et postdoctorales



Ce mémoire intitulé :

**Étude sur la reconnaissance de l'ubiquitine par les domaines de transactivation  
acides des activateurs de transcription**

Présenté par :

Mathieu Lussier-Price

a été évalué par un jury composé des personnes suivantes :

Luc DesGroseillers, président-rapporteur

James G. Omichinski, directeur de recherche

André Tremblay, membre du jury







## Résumé

Les domaines de transactivation (TAD) acides sont présents dans plusieurs protéines oncogéniques, virales et dans des facteurs de différenciation de cellules souches. Ces domaines acides contrôlent la transcription à travers une myriade d'interactions avec divers partenaires ce qui provoque l'activation de la transcription ou leur propre élimination. Cependant, dans la dernière décennie, de plus en plus de recherches ont démontré que les TAD possédaient un sous-domaine activation/dégradation (DAD) responsable pour une fonction d'activation de la transcription dépendante de la dégradation de la protéine. Un tel phénomène peut être accompli par plusieurs moyens tels que des modifications post-traductionnelles, l'association à des cofacteurs ou la formation d'un réseau d'interaction complexe en chaînes. Or, aucune preuve concrète n'a pu clairement démontrer le fonctionnement de la dépendance paradoxale entre ces deux fonctions sur un activateur de transcription.

Le DAD, a été observé dans plusieurs facteurs de transcription incluant la protéine suppresseur de tumeur p53 et le facteur de différenciation érythrocyte EKLF. Un aspect particulier des DAD est que la composition de leur séquence d'acide aminé est fortement similaire à celle des domaines de liaison à l'ubiquitine (UBD) qui jouent un rôle clé dans le contrôle de la transcription à travers leur interaction non-covalente avec l'ubiquitine.

Ainsi, dans ce mémoire, nous avons étudié la possibilité que les TAD acides soient capables d'agir comme UBD pour réguler leur fonction paradoxale à travers des interactions non-covalentes avec l'ubiquitine. L'analyse est faite en utilisant la résonance magnétique nucléaire (RMN) ainsi qu'avec des essais fonctionnels de dégradation. En somme, cette étude amène une plus grande compréhension des protéines impliquées dans le contrôle des TAD et caractérise le tout premier exemple de TAD capable d'interagir avec l'ubiquitine.

**Mots-clés :** Activateur de transcription acide, domaine d'activation et de dégradation (DAD), domaine de liaison à l'ubiquitine (UBD), Erythroïde Kruppel-Like-Factor (EKLF), protéine suppresseuse de tumeur p53, interaction protéine-protéine (IPP), résonance magnétique nucléaire (RMN).





## Abstract

Acidic transactivating domains have been shown to be potential targets for a number of different therapies but their dynamic nature and their ability to bind many interacting partners has made it difficult to fully understand their functioning mechanisms. What we do know about these domains is that they readily control transcription through a myriad of interactions capable of either activating specific aspects of their function or simply, signal for their own demise. Within the acidic TADs lies an unusual degradation/activation domain (DAD) capable of activating transcription at the cost of its degradation. In other words, DAD transcriptional activation is dependent on the degradation of the protein. Such a phenomenon could be explained by a wide variety of hypotheses like the play of post-translational modifications, co-factors, or maybe just a really sophisticated time scaled network of interactions. However, no concrete explanation of how this dual dependent functioning domain works has yet to surface.

The DAD has been observed within acidic TADs of several transcription factors including the tumor suppressor p53 and the red blood cell differentiation factor EKLF. Interestingly though, the amino acid sequence composition of DADs share a strong similarity with several types of sequences from domains that bind ubiquitin (UBDs). These domains have been shown in the past to, in addition to their role in degradation, play a key role in regulating transcription through non-covalent interaction with ubiquitin. Hence, in this project, we investigated whether acidic TADs had the ability to function as UBDs and form non-covalent interactions with ubiquitin and also to determine the functional significance of this interaction in regards to the dual function of acidic TADs.

**Keywords:** Acidic transactivating domains, Degradation activation domain (DAD), Ubiquitin binding domain (UBD), Ubiquitin (UBI), Erythroide Kruppel-Like-Factor (EKLF), Tumor suppressor protein p53, protein-protein interaction (PPI), nuclear magnetic resonance (NMR).



# Table des matières

|   |             |
|---|-------------|
| <b>Résumé.....</b>  | <b>I</b>    |
| <b>Abstract.....</b>  | <b>III</b>  |
| <b>Liste des tableaux .....</b>   | <b>VIII</b> |
| <b>Liste des figures .....</b>  | <b>IX</b>   |
| <b>Liste des abréviations .....</b>   | <b>XI</b>   |
| <b>Remerciements.....</b>   | <b>XV</b>   |
| <b>1. Introduction.....</b>   | <b>1</b>    |
| 1.1 Les activateurs de transcription .....  | 1           |
| 1.1.1 <i>La transcription et le rôle des activateurs de transcription. ....</i>   | <i>1</i>    |
| 1.1.2 <i>Les domaines des activateurs de transcription .....</i>  | <i>2</i>    |
| 1.1.3 <i>Le TAD acide d'EKLf et la régulation de l'érythropoïèse .....</i>  | <i>3</i>    |
| 1.1.4 <i>Le TAD acide de p53 et le rôle du gardien du génome.....</i>   | <i>5</i>    |
| 1.1.5 <i>Similitudes des TAD d'EKLf et de p53.....</i>  | <i>8</i>    |
| 1.1.6 <i>Fonction paradoxale des TAD acides .....</i>   | <i>11</i>   |
| 1.2 Le système de l'ubiquitine.....   | 13          |
| 1.2.1 <i>L'ubiquitination .....</i>   | <i>13</i>   |
| 1.2.2 <i>Les différents signaux de l'ubiquitine.....</i>  | <i>14</i>   |
| 1.2.3 <i>La reconnaissance de l'ubiquitine et les domaines de liaison à l'UBI. ....</i>                                 | <i>16</i>   |
| 1.2.4 <i>La surface d'interaction L8-I44-V70 de l'ubiquitine .....</i>  | <i>17</i>   |
| 1.3 Le système d'ubiquitine dans la régulation des activateurs.....   | 19          |
| 1.3.1 <i>Rôle protéolytique .....</i>   | <i>19</i>   |
| 1.3.2 <i>Rôle non-protéolytique.....</i>  | <i>20</i>   |
| 1.3.3 <i>Le système de l'ubiquitine et les deux modèles du mécanisme de la fonction paradoxale des activateurs.....</i> | <i>21</i>   |
| 1.4 Hypothèse et objectifs .....  | 23          |
| 1.5 Techniques utilisées .....  | 28          |
| 1.5.1 <i>Production et purification de protéines pour la RMN .....</i>  | <i>29</i>   |

|         |   |    |
|---------|---|----|
| 1.5.2   | <i>Le titrage RMN pour détecter et caractériser les interactions entre protéines.</i> | 30 |
| 1.5.3   | <i>Les PDC pour quantifier la force d'interaction. ....</i>                           | 31 |
| 1.5.4   | <i>Déterminer la structure d'une interaction entre protéines à l'aide de la RMN</i>   | 33 |
| 1.5.4.1 | <i>Collection de contraintes structurales ....</i>                                    | 34 |
| 1.5.4.2 | <i>Calcul et validation de la structure ....</i>                                      | 35 |
| 1.5.5   | <i>Analyse des structures à l'aide d'études fonctionnelles. ....</i>                  | 36 |

## **2. Caractérisation d'une interaction entre le TAD d'EKLf et l'ubiquitine ..39**

|      |  |    |
|------|--|----|
| 2.1. | Mise en contexte de l'article.....   | 41 |
| 2.2. | Contribution des auteurs .....   | 42 |
| 2.3. | The acidic TAD of EKLf directly binds ubiquitin.....                             | 43 |
|      | <i>SUMMARY.....</i>  | 44 |
|      | <i>INTRODUCTION.....</i>   | 45 |
|      | <i>RESULTS .....</i>   | 48 |
|      | <i>EKLfTAD1 Binds Non-covalently to UBI .....</i>                                | 48 |
|      | <i>UBI Interacts Non-covalently with EKLfTAD1 .....</i>                          | 49 |
|      | <i>NMR Structure Determination of the UBI-EKLfTAD1 Complex.....</i>              | 50 |
|      | <i>Overall Structure of the EKLfTAD1-UBI Complex.....</i>                        | 50 |
|      | <i>Mutational Studies of the EKLfTAD1 Interaction with UBI .....</i>             | 52 |
|      | <i>The Role of the UBI-interacting region of EKLfTAD1 in EKLf Stability.....</i> | 53 |
|      | <i>The Activation-Degradation Region of SREBP1a Binds UBI .....</i>              | 53 |
|      | <i>DISCUSSION.....</i>   | 54 |
|      | <i>EXPERIMENTAL PROCEDURES .....</i>   | 57 |
|      | <i>Cloning of EKLf-RLuc Constructs. ....</i>                                     | 57 |
|      | <i>EKLf Stability Assay. ....</i>  | 58 |
|      | <i>Statistical Analysis.....</i>   | 58 |
|      | <i>Cloning, Expression and Purification of Recombinant Proteins. ....</i>        | 58 |
|      | <i>NMR Samples. ....</i>   | 59 |
|      | <i>NMR Spectroscopy. ....</i>  | 59 |

|   |           |
|---|-----------|
| <i>Structure Calculations.</i> .....                                      | 60        |
| <i>ACKNOWLEDGEMENTS</i> .....   | 60        |
| <i>FIGURES</i> .....  | 61        |
| <i>REFERENCES</i> .....   | 68        |
| <i>SUPPLEMENTARY EXPERIMENTAL PROCEDURES</i> .....                        | 73        |
| <i>SUPPLEMENTAL REFERENCE</i> .....                                       | 73        |
| <i>SUPPLEMENTARY TABLE</i> .....  | 74        |
| <i>SUPPLEMENTARY FIGURES</i> .....  | 75        |
| <b>3. Discussion</b> .....  | <b>81</b> |
| 3.1 Retour sur les résultats.....   | 83        |
| 3.2. Projets futurs.....  | 88        |
| 3.2.1. La liaison à l'UBI et le modèle d'activation négative .....        | 88        |
| 3.2.2. La liaison à l'UBI et les TAD viraux .....                         | 92        |
| 3.2.3. La liaison à l'UBI, la E3 ligase et les TAD de p53 et d'EKLf ..... | 94        |
| 3.2.3.1. p53, MDM2 et UBI .....   | 94        |
| 3.2.3.2. EKLf, UBI et une E3 inconnue .....                               | 97        |
| 3.3. Conclusion générale .....  | 98        |
| <b>Bibliographie</b> .....  | <b>i</b>  |

## Liste des tableaux

|  |    |
|--|----|
| Tableau I. NMR and refinement statistics for EKLFTAD1-UBI complex <sup>a</sup> ..... | 74 |
|--|----|

## Liste des figures

### Chapitre 1.

|   |    |
|---|----|
| Figure 1.1. Le rôle des activateurs de transcription dans la transcription.....     | 2  |
| Figure 1.2. Le rôle d'EKLf dans l'érythropoïèse. ....                               | 4  |
| Figure 1.3. Domaines fonctionnels d'EKLf.....                                       | 5  |
| Figure 1.4. Domaines fonctionnels p53. ....   | 7  |
| Figure 1.5. Régulation des niveaux et de l'activité de p53 par MDM2 et MDMX. ....   | 8  |
| Figure 1.6. Comparaison entre les TAD de p53 et d'EKLf.....                         | 10 |
| Figure 1.7. Les TAD et les « degron » des activateurs de transcription.....         | 12 |
| Figure 1.8. Les différentes étapes de l'ubiquitination. ....                        | 14 |
| Figure 1.9. Les différentes lysines présentes chez l'ubiquitine. ....               | 15 |
| Figure 1.10. Les différents signaux de l'ubiquitination.....                        | 15 |
| Figure 1.11. La surface d'interaction L8-I44-V70 de l'ubiquitine.....               | 18 |
| Figure 1.12. Le protéasome 26S. ....  | 19 |
| Figure 1.13. Les deux modèles du mécanisme de contrôle des domaines paradoxaux..... | 22 |
| Figure 1.14. Comparaison entre les séquences d'UIM, MIU et DAD. ....                | 24 |
| Figure 1.15. Comparaison entre les domaines UIM et MIU. ....                        | 25 |
| Figure 1.16. Expérience de titrage par RMN-2D. ....                                 | 31 |
| Figure 1.17. Simulation de courbes pour quantifier le $K_D$ à partir des PDC.....   | 33 |
| Figure 1.18. Schéma du principe des spectres NOESY.....                             | 35 |

### Chapitre 2.

|  |    |
|--|----|
| Figure 2. 1. The TAD of EKLf binds to UBI.....   | 61 |
| Figure 2. 2. EKLFTAD1 binds the Leu8-Ile44-Val70 hydrophobic patch of UBI. ....  | 62 |
| Figure 2. 3. Structure of the EKLFTAD1-UBI complex. ....   | 63 |
| Figure 2. 4. Structural details of the EKLFTAD1-UBI interface. ....  | 65 |
| Figure 2. 5. The minimal TAD of EKLf functions as a “degron” in cells. ....  | 65 |
| Figure 2. 6. The acidic ADR of SREBP-1a binds to UBI.....  | 66 |
| Figure 2. 7. Sequence and structure comparison of the UBI-interacting helices from EKLFTAD1, Rabex5 and Vps27.....                           | 67 |
| Supplementary Figure 2.S 1 related to Figure 2. 1. EKLFTAD2 binds weakly to UBI. ....  | 75 |
| Supplementary Figure 2.S 2 related to Figure 2. 4. Positively charged residues of UBI help to position the negatively charged EKLFTAD1. .... | 76 |
| Supplementary Figure 2.S 3 related to Figure 2. 4. Phe27, Leu28 and Trp31 from EKLFTAD1 are crucial for binding to UBI. ....                 | 77 |
| Supplementary Figure 2.S 4 related to Figure 2. 4. Ile44 of UBI is a crucial for EKLFTAD1 binding to UBI.....                                | 78 |

|   |    |
|---|----|
| Supplementary Figure 2.S 5 related to Figure 2. 6. p53TAD1, Gal4TAD and VP16NTAD do not bind UBI with significant affinity. ....                                  | 79 |
| Supplementary Figure 2.S 6 related to Figure 2. 6. Sequence Homology between EKLFTAD1 and other acidic TADs that function as activation-degradation regions. .... | 79 |

### Chapitre 3.

|  |    |
|--|----|
| Figure 3. 1. Modèle proposé de la régulation des domaines paradoxaux d'EKLf et de SREBP1a..... | 85 |
| Figure 3. 2. Le TAD de p53 et sa capacité à lier l'UBI. ....                                   | 90 |
| Figure 3. 3. Effet de la mutation p53_W53S sur la formation du complexe avec l'UBI. ....       | 91 |
| Figure 3. 4. Le TADC de VP16 et sa capacité à lier l'UBI. ....                                 | 93 |
| Figure 3. 5. Complexe ternaire entre p53, UBI et MDM2. ....                                    | 96 |



## Liste des abréviations

<sup>13</sup>C : Carbone 13

<sup>15</sup>N : Azote 15

ADN : Acide désoxyribonucléique (deoxyribonucleic acid, DNA)

ARF: Protéine suppresseur de tumeur à cadre de lecture alternatif (Alternate reading frame tumor suppressor protein)

ARN : Acide ribonucléique (ribonucleic acid, RNA)

ARNm : ARN messenger (messenger RNA, mRNA)

AT : Activateur de transcription (transcriptional activator)

ATPase : Enzyme qui hydrolyse ou synthétise les molécules d'adénosine-triphosphate

CBP : Protéine de liaison à CREB (CREB binding protein)

cFos: Oncogène cellulaire Fos (cellular oncogene Fos)

CLHP : Chromatographie en phase liquide à haute performance (high performance liquid chromatography, HPLC)

CSH: Cellules souches hématopoïétiques

CUE : Domaine de couplage de l'ubiquitine à la dégradation de protéines du réticulum endoplasmique (Coupling of ubiquitin to ER degradation)

DAD : Domaine d'activation et de dégradation (degradation activation domain)

DBD : domaine de liaison à l'ADN (DNA-binding domain)

DUB: enzyme de-ubiquitinase (deubiquitinating enzyme)

E1: Enzyme activatrice de l'ubiquitine (ubiquitin activating enzyme)

E2: Enzyme de conjugaison de l'ubiquitine (ubiquitin conjugating enzyme)

E3: Enzyme ligase de l'ubiquitine (ubiquitin ligase)

EBNA : l'antigène nucléaire du virus d'Epstein-Barr (EBV nuclear antigen)

EKLF : Facteur de transcription érythrocyte Kruppel-like (Erythroide kruppel -like factor)

EKLFTAD1 : Résidus 1-40 du facteur de transcription EKLF

EKLFTAD2 : Résidus 51-90 du facteur de transcription EKLF

Fra-1: Antigène 1 relié à Fos (Fos-related antigen 1)

FTIIB : facteur général de transcription II B (transcription initiation factor IIB)

FTIID : facteur général de transcription II D (transcription initiation factor IID)

FTIIE : Facteur général de transcription II E (transcription initiation factor IIE)

FTIIF : Facteur général de transcription II F (transcription initiation factor IIF)

FTIIH : Facteur général de transcription II H (transcription initiation factor IIH)

Gal-4 : Protéine régulatrice 4 chez la levure (Yeast Regulatory protein GAL4)

GST : Glutathion-S-Transférase (glutathion-S-Transférase)

H2A : Histone A

H2B : Histone B

HPV : Virus de l'herpès humain (human herpes virus)

HSQC : Cohérence simple quantum hétéronucléaire (Heteronuclear single quantum coherence)

hSWI/SNF : Interrupteur humain/sucrose non fermentable (Human SWItch/Sucrose Non Fermentable)

IPP: Interaction protéine/protéine (protein/protein interaction, PPI)

$K_D$  : Constante de dissociation (dissociation constant)

MDM2 : Protéine murine double minute 2 (mouse double minute 2 homolog)

MDM4 ou MDMX: Protéine murine double minute 4 ou X (murine double minute 4 or X homolog)

MIU : motifs interagissant avec l'ubiquitine (Motif interacting with ubiquitin)

NLS : Signal de localisation nucléaire (nuclear localisation signal)

NOE : Effet nucléaire Overhauser (Nuclear Overhauser effect)

NOESY : Spectroscopie à effet nucléaire Overhauser (Nuclear Overhauser effect spectroscopy)

NSs : Protéine non-structurale de segment S du virus de la vallée du Rift (Non-structural protein NSs)

p14ARF: Protéine 14 cadre de lecture alternatif (protein 14 alternate reading frame)

- p300 : Protéine de liaison à la protéine E1A (E1A binding protein)
- p53 : Protéine de suppression de tumeur 53 (tumor suppressor protein p53)
- p53TAD1 : Résidus 1-40 du facteur de transcription p53
- p53TAD2 : Résidus 40-73 du facteur de transcription p53
- p62 : Sous-unité 1 du FTIIH chez l'humain (subunit of human FTIIH)
- PH : Homologie à la pleckstrin (pleckstrin homology)
- PI: Point isoélectrique (isoelectric point, IP)
- PIC : complexe de préinitiation (preinitiation complex)
- pol II : ARN polymérase II (RNA polymerase II)
- ppm : Partie par million (parts-per-million)
- RMN : Résonance magnétique nucléaire (nuclear magnetic resonance, NMR)
- S5a: Sous-unité homologue non-ATPase régulatrice 4 (multiubiquitin binding protein S5a)
- SAX: Diffusion de rayons X aux petits angles (small angle X-ray scattering)
- SREBP1a : Protéine 1a liant les éléments régulateur du stérol (Sterol regulatory element-binding protein 1)
- TAD : Domaine de transactivation (transactivation domain)
- TBP : Protéine liant TATA (TATA-binding protein)
- Tet : Tétramérisation
- Tfb1 : Sous-unité 1 du facteur général de transcription IIH chez la levure (RNA polymerase II transcription factor B subunit 1)
- UBD : Domaine de liaison à l'ubiquitine (ubiquitin binding domain)
- UBI : Ubiquitine (ubiquitin)
- UIM : Ubiquitine interactant (Ubiquitin interacting motif)
- VP16 : Protéine virale 16 (viral protein 16)
- Vps27: Protéine vacuolaire 27 associée au triage (Vacuolar sorting-associated protein 27)
- $\delta$  : Déplacement chimique (chemical shift)



## Remerciements

Je tiens à remercier tout d'abord mon directeur de maîtrise, le Dr James G. Omichinski pour tout ce qu'il m'a appris durant ces dernières années. Merci pour votre soutien et pour l'immense générosité dont vous faites preuve pour accompagner tous vos étudiants, dont j'ai eu la chance de faire partie. Votre charisme contagieux et le savoir dont vous m'avez transféré m'ont donné goût non seulement à continuer dans le domaine de biochimie, mais aussi à me dépasser. Ce fut un plaisir de venir travailler dans votre laboratoire durant ces trois années et j'espère pouvoir mettre à profit dans le futur tout ce que vous m'avez appris.

Je voudrais remercier les Dr Luc DesGroseillers et Dr André Tremblay d'avoir accepté d'être, respectivement, président-rapporteur et membre du jury pour mon mémoire. Merci au jury, qui évaluera ce mémoire pour son temps et son attention accordés à mes travaux.

Merci à tous les membres des laboratoires Ominchinski et Legault, avec qui j'ai eu la chance de travailler. Vous avez tous contribué à créer une belle ambiance de travail. Je tiens tout particulièrement à remercier Éric Bonneau, Lauriane Lecoq et Philippe Chabot pour la relecture de ce manuscrit. Merci Caroline, Luca et Julien pour leur disponibilité à discuter ou répondre à mes nombreuses questions de RMN ou à enregistrer des spectres avec moi, Laurent et Haytham pour m'avoir initié au monde de cristallographie, Xavier pour m'avoir montré le côté *in vivo* de la biochimie, Geneviève pour m'avoir aidé avec mes expériences biomoléculaires, Philippe pour qui j'espère n'a pas trop souffert et sans oublier Éric, pour les nombreuses discussions lors de nos pauses café. Vous avez tous participé à créer de belles années au cours desquelles j'ai pu apprendre énormément, mais aussi faire de bons amis.

Finalement, je voudrais remercier infiniment ma famille. Mes parents, pour leur soutien tout au long de mes études, Kumo et Nikki pour leur soutien moral durant les temps difficiles. Et enfin, Jenny, pour ton soutien et ne jamais avoir douté de mes capacités...



## **1. Introduction**





## **1.1 Les activateurs de transcription**

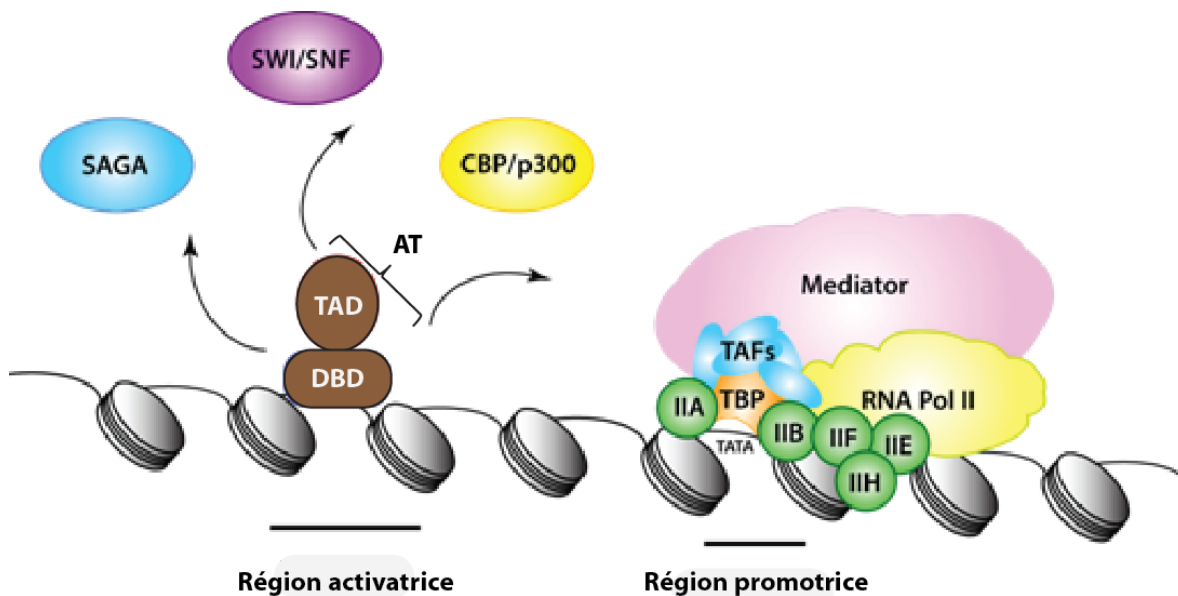
Les activateurs de transcription (AT) sont des éléments essentiels à la survie et au maintien des cellules. Ces protéines orchestrent l'enchaînement des mécanismes qui permettent le bon fonctionnement de la cellule en choisissant les gènes à exprimer en temps et en lieu appropriés. Le rôle des AT peut être observé dans une grande variété de type de régulation comme le développement embryonnaire, l'hydratation des reins ou encore en évitant que la cellule active accidentellement des mécanismes non-réversibles comme son apoptose [1]. Plusieurs maladies découlent du fait que la transcription des gènes est dérégulée. D'après plusieurs études, les AT peuvent être ciblés pour remettre à niveau la transcription des gènes associés à ces maladies [2, 3]. Néanmoins, le potentiel d'utilisation des AT est encore limité par un manque de compréhension fonctionnelle et structurale sur leur mécanisme de contrôle de la transcription.

### ***1.1.1 La transcription et le rôle des activateurs de transcription.***

Chez les eucaryotes, la transcription est le processus par lequel les différents types d'ARN sont synthétisés. Parmi ceux-ci, les ARN messagers (ARNm) sont ceux qui constituent la matrice de production des protéines. Il y a trois grandes étapes lors de la transcription de l'ADN en ARNm : l'initiation (synthèse des premiers nucléotides), l'élongation (l'allongement de l'ARNm) et la terminaison (relâchement de l'ARNm). Ces trois processus sont tous effectués par l'ARN polymérase II (ARN pol II). Or, pour initier la première étape de la transcription, l'ARN pol II ne peut fonctionner sans l'aide des facteurs de transcription (FT) généraux. Ensemble, l'ARN pol II et les FT généraux (FTIIA, FTIIB, FTIID, FTIIE, FTIIF et FTIIH) forment le complexe de préinitiation (PIC) qui constitue la machinerie transcriptionnelle de base pour la majorité des gènes (pour une revue plus détaillée sur la transcription, voir [4]). Cependant, face à un stimulus, ce sont les AT qui, aidés de co-activateurs, dicteront la variation du taux de transcription d'un gène ou d'un groupe de gènes spécifiques pour répondre aux besoins de la cellule.

### 1.1.2 Les domaines des activateurs de transcription

Les AT comprennent à la base deux domaines fonctionnels interchangeables : un domaine de liaison à l'ADN (DBD) et un domaine d'activation de la transcription (TAD). Les DBD permettent aux activateurs de lier une séquence spécifique d'ADN (région activatrice). Plusieurs motifs structuraux sont connus pour lier l'ADN tels que les domaines à doigts de zinc, les domaines hélice-tour-hélice ou encore les glissières à leucine [5]. Lorsque liés à l'ADN, les AT utilisent leur domaine TAD pour stimuler la transcription. Les TAD activent la transcription via leur interaction avec des partenaires protéiques communs. Parmi ceux-ci, on retrouve des facteurs généraux de la transcription (TBP, FTIIB, TAF31, TAF80, TAF9, FTIIH) [6-9], des histones acétyltransferases (le complexe SAGA), des co-activateurs (CBP/p300), des composants de remodelage du nucléosome (SWI/SNF : SWItching mating type/Sucrose Non Fermenting) et des composants du complexe médiateur (MED15, MED25) (**Figure 1.1**) [10, 11].



**Figure 1.1. Le rôle des activateurs de transcription dans la transcription.**

Représentation schématique des facteurs qui initient la transcription et le rôle exercé par un activateur de la transcription. À droite; les éléments de la machinerie transcriptionnelle de base (le PIC, ses sous-unités et l'ARN pol II). À gauche; les différentes sous-unités d'un AT (en brun) lié à l'ADN (en gris) et ses différents partenaires d'interaction potentiels lors de la transcription (représenté par les flèches).

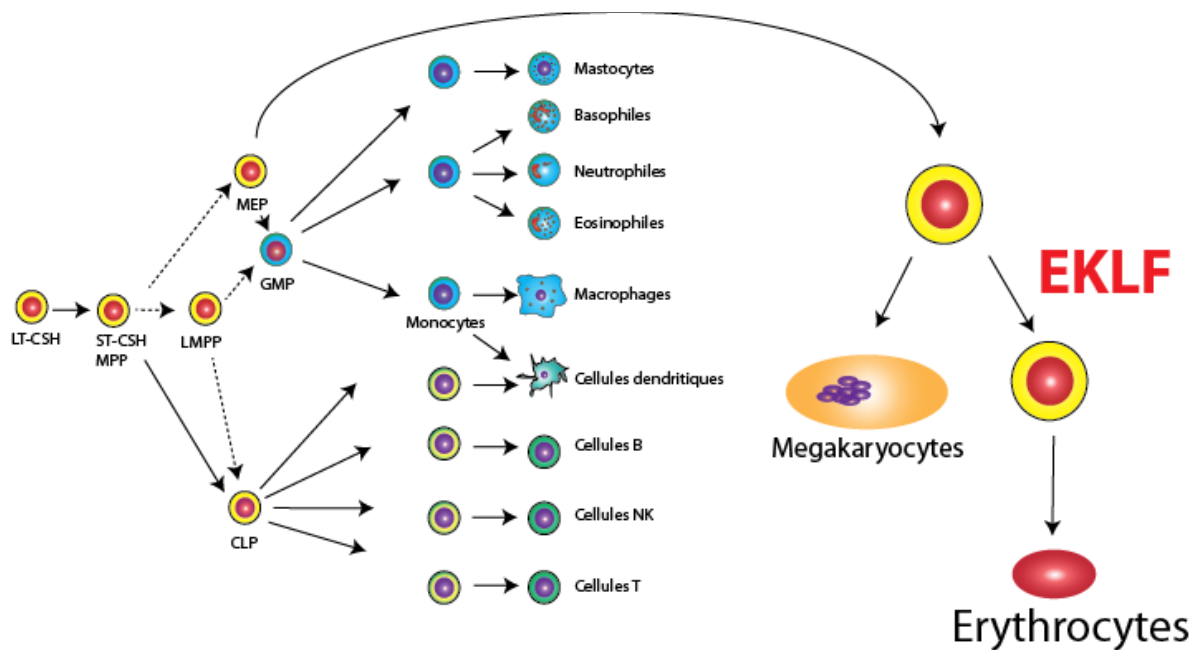
Les TAD sont traditionnellement classés en fonction de leur composition en acides aminés [12]. On distingue les domaines riches en résidus glutamine, riches en proline, riches en sérine et thréonine et enfin, riches en résidus acides. La classe des TAD acides se distingue des autres puisqu'elle est présente dans plusieurs protéines pro-oncogéniques (p53, myc, Fox1), dans des protéines de virus potentiellement cancérogènes (VP16, EBNA2) et dans des facteurs de différenciation des cellules souches (EKLF, p65 et E2F1) [13-16]. Une des caractéristiques particulières de ces TAD acides est qu'ils sont des domaines intrinsèquement désordonnés (sans structure stable) sous forme libre. Cependant, ils peuvent prendre diverses formes structurées lorsqu'ils interagissent avec des protéines cibles. Ceci leur offre des avantages fonctionnels tels que la possibilité d'interagir avec plusieurs partenaires à cause de leur capacité à adapter leur structure en fonction des sites de liaison [17]. Ainsi, la souplesse des TAD acides leur permet de jouer un rôle clé dans plusieurs AT. Parmi les AT possédant des TAD acides se trouve le facteur de différenciation érythropoïétique EKLF qui contrôle la maturation des érythrocytes et aussi la protéine suppresseur de tumeur p53.

### **1.1.3      *Le TAD acide d'EKLF et la régulation de l'érythropoïèse***

L'érythropoïèse définit l'ensemble des processus de production des érythrocytes dans la moelle osseuse à partir de cellules souches hématopoïétiques multipotentes (CSH). Chez l'humain, la durée de vie des globules rouges est comprise entre 110 et 120 jours. Les globules rouges ont un rôle capital dans l'organisme puisqu'ils sont chargés de transporter l'oxygène des poumons vers les tissus. Plusieurs maladies découlent d'une dérégulation de ce système incluant l'anémie [18] ainsi que des maladies plus rares comme certains cas de porphyrie [19]. Depuis plusieurs années, des études ont permis de mieux comprendre l'érythropoïèse et les facteurs de transcription qui permettent le passage entre chaque lignée cellulaire.

Parmi ces facteurs, EKLF est essentiel pour guider le passage du stade fœtal au stade adulte des globules rouges. Plus précisément, EKLF active les gènes permettant à la  $\beta$ -globine de remplacer la  $\gamma$ -globine empêchant ainsi la formation de mégacaryocytes (une

lignée différente de l'érythropoïèse) (**Figure 1.2**) [20]. Le rôle crucial d'EKLK est illustré par le fait que la délétion d'EKLK, chez les souris, provoque la mort de celle-ci par anémie au cours de l'érythropoïèse foétale [21].

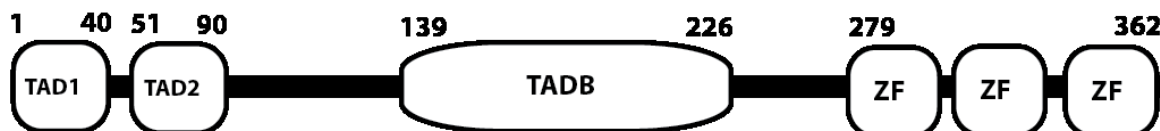


**Figure 1.2. Le rôle d'EKLK dans l'érythropoïèse.**

Les différentes étapes de la différenciation en érythrocytes à partir de cellules pluripotentes et l'implication du facteur de transcription EKLK. LT-CSH, cellules souches hématopoïétiques long-terme; ST-CSH, cellules souches hématopoïétiques court-terme; MPP, progéniture multipotente; LMPP, progéniture multipotente primé-lymphoïde; MEP, progéniture mégacaryocyte/érythroïde; GMP, progéniture granulocyte/macrophage; CLP, progéniture lymphoïde commune. Figure adaptée de [22].

EKLK est un membre de la famille des quinze protéines « Krüppel like factors » (KLF 1-15) qui tiennent leur nom grâce à leur ressemblance aux protéines encodées par le gène du même nom. Ces protéines sont caractérisées par la présence de trois domaines doigts de zinc Cys2His2 à leur extrémité carboxyle-terminale qui sont utilisés pour lier des régions promotrice sur l'ADN comme les sites CACCC ou de type boîte GT. La présence de trois domaines doigts de zinc permet une meilleure spécificité pour les gènes régulés par EKLK. Contrairement aux autres membres de la famille des KLF, EKLK (aussi appelé KLF1) est exprimé uniquement dans les érythrocytes. EKLK lie spécifiquement le site promoteur du gène de  $\beta$ -globine (**Figure 1.3**) [20]. Des mutations au niveau des doigts de zinc d'EKLK en C-terminal ou dans la séquence promotrice du gène résultent en une diminution de l'expression du gène  $\beta$ -globine, engendrant des maladies liées aux globules rouges

comme la  $\beta$ -thalassémie, l'anémie néonatale ou la persistance d'hémoglobine fœtale héréditaire [21, 23-26]. À son extrémité N-terminale, EKLF contient un domaine d'activation qui inclue un TAD acide localisé dans les premiers 90 résidus [27]. Ce TAD d'EKLF peut être divisé en deux parties indépendantes : le domaine EKLTAD1 (résidus 1-40) et le domaine EKLTAD2 (résidus 51-90) qui sont indépendamment capables d'activer la transcription [27].



**Figure 1.3. Domaines fonctionnels d'EKLF.**

Diagramme représentant les domaines d'EKLF où TAD1, TAD2 et TADB sont les domaines de transactivation et DBD formé de trois doigts de zinc (ZF). Les nombres représentent la position dans la séquence.

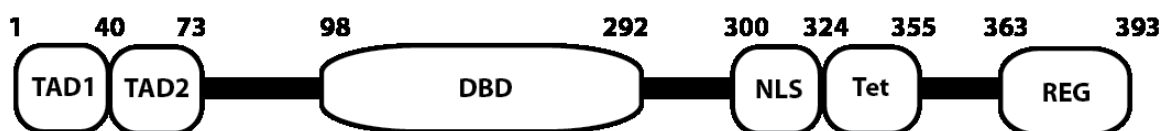
EKLTAD1 et EKLTAD2 sont aussi connus pour posséder des régions appelées PEST qui sont des régions riches en résidus prolines, acides glutamiques, sérines et thréonines. Les régions PEST sont souvent associées à la dégradation des protéines par le système de l'ubiquitine [28]. En effet, la délétion des régions TAD en N-terminal, spécialement la région EKLTAD1, ralentissait de manière significative la dégradation d'EKLF sans affecter l'ubiquitination de celle-ci [29]. Ces résultats suggèrent que les domaines TAD d'EKLF possèdent des fonctions d'activation et de dégradation qui se chevauchent et qui sont critiques pour le rôle d'EKLF dans le développement de globules rouges.

#### **1.1.4 Le TAD acide de p53 et le rôle du gardien du génome**

L'AT p53 est une protéine qui est exprimée dans tous les types de cellules humaines, mais se trouve cependant à de basses concentrations cellulaires sous des conditions normales [30, 31]. Elle est activée quand la cellule est exposée à différents stress comme des dommages à l'ADN et cette activation peut mener à l'arrêt de la croissance cellulaire, à l'activation des mécanismes de réparation de l'ADN, à la sénescence ou encore à l'apoptose [31-34]. La fonction de p53 est de maintenir l'intégrité du génome. On réfère souvent à cette protéine comme étant le gardien du génome. Si la fonction de p53 est altérée à travers des mutations, des délétions au niveau du gène ou

encore à travers une dérégulation quelconque, l'apparition de cancer est quasiment inévitable. Ainsi, il n'est pas étonnant de voir que p53 est muté dans plus de 50 % de plusieurs types de cancers [32]. Cependant, le rôle de p53 peut s'étendre dans plusieurs autres domaines incluant le contrôle du vieillissement ou dans des pathologies neuro-dégénératives telles que Alzheimer, Parkinson et Huntington [35-38].

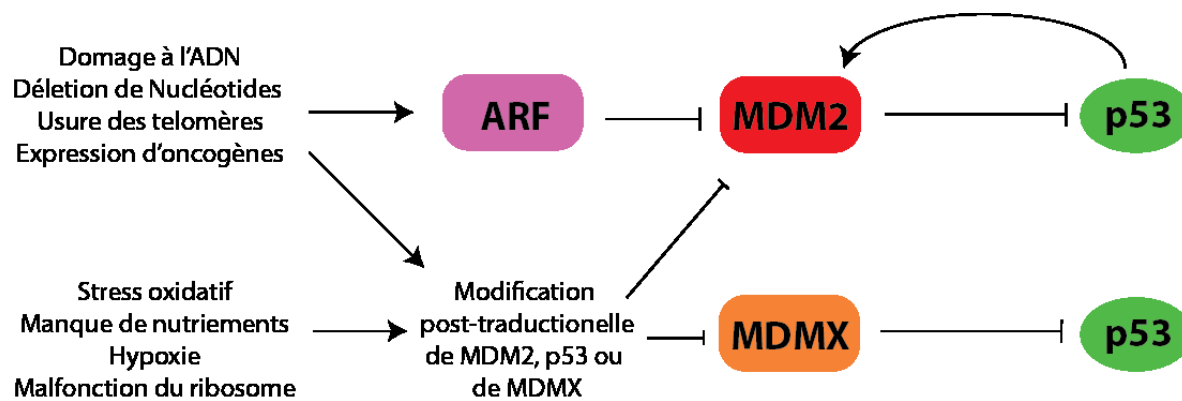
La protéine p53 est formée de plusieurs domaines modulaires (**Figure 1.4**). À son extrémité N-terminale, p53 contient deux domaines de transactivation acides, p53TAD1 et p53TAD2, qui s'étendent des résidus 1 à 40 et 40 à 73 respectivement [39, 40]. Ces domaines peuvent indépendamment activer la transcription des gènes ciblés par le biais d'interactions avec des éléments de la machinerie transcriptionnelle [41]. La redondance fonctionnelle de ces deux TAD agirait comme alternative dans le cas où l'un serait défectueux. Ceci pourrait expliquer pourquoi on ne trouve presque pas de mutations dans ces domaines chez des patients atteints de cancer. La majorité des cancers qui affectent cet activateur possèdent des mutations plutôt au niveau du domaine DBD de p53 (résidus 98 à 292). En effet, la mutation d'un seul acide aminé dans ce domaine peut non seulement abolir sa capacité à lier l'ADN, mais aussi son activité transcriptionnelle. La protéine p53 possède aussi d'autres domaines comme le domaine de signal de localisation cellulaire NLS (résidus 300 à 324) et de tétramérisation Tet (résidus 324 à 355) qui servent de régulation additionnelle pour p53. De plus, p53 possède en partie C-terminale un domaine chargé négativement (résidus 363 à 393) qui est la cible de plusieurs modifications post-traductionnelles. Ce dernier est capable de lier l'ADN de manière non-spécifique ce qui servirait d'interrupteur pour aider le DBD à lier les séquences appropriées face à différents stimuli. En effet, des mutations ponctuelles ou délétions au niveau de cette partie C-terminale empêche la liaison de p53 à des régions promotrices de l'ADN [42].



**Figure 1.4. Domaines fonctionnels p53.**

Diagramme représentant les domaines de p53 où TAD : domaine de transactivation, DBD : domaine de liaison à l'ADN, Tet : domaine de tétramérisation, C-term : domaine C-terminale régulateur.

Plusieurs mécanismes ont été identifiés pour réguler l'activité de p53 à partir de son TAD. Les enzymes ligases d'ubiquitine (E3), particulièrement la ligase murine double minute 2 (MDM2), agissent comme régulateurs négatifs de p53 en empêchant son accumulation en cas normal [43]. MDM2 lie p53TAD1 et, non seulement régule son accumulation, mais aussi inhibe sa transactivation en bloquant le recrutement d'éléments essentiels pour la transcription. L'activation de p53 est possible soit (1) lorsque MDM2 et p53 sont modifiés post-traductionnellement par une phosphorylation suite à un stimuli; (2) lorsque MDM2 est séquestré par l'induction oncogénique de la protéine de suppression de tumeurs p14ARF ou (3) lorsque MDM2 interagit avec des protéines ribosomales induites par des stress nucléolaires [44-46]. Cependant, plusieurs autres mécanismes sont mis en place pour que p53 soit bien régulé. Entre autres, il a été démontré qu'une protéine de la même famille que MDM2; murine double minute X (MDMX) (aussi appelé MDM4) a un effet similaire de liaison qui bloque l'activité de p53 [47]. Malgré son incapacité à agir en tant que E3 ligase, MDMX agit comme enzyme de soutien dans le cas où MDM2 serait défectueux en occupant le p53TAD1 (**figure 1.5**) [48, 49].



**Figure 1.5. Régulation des niveaux et de l'activité de p53 par MDM2 et MDMX.**

Niveaux et activité de p53 régulés par la E3 ligase MDM2 via ubiquitination, blocage de son domaine TAD et par MDMX en bloquant aussi son TAD. MDM2 est régulé par des modifications post-traductionnelles, ARF ou par des protéines ribosomales suite à différents stress cellulaires. Figure adaptée de [50].

#### **1.1.5 Similitudes des TAD d'EKLf et de p53**

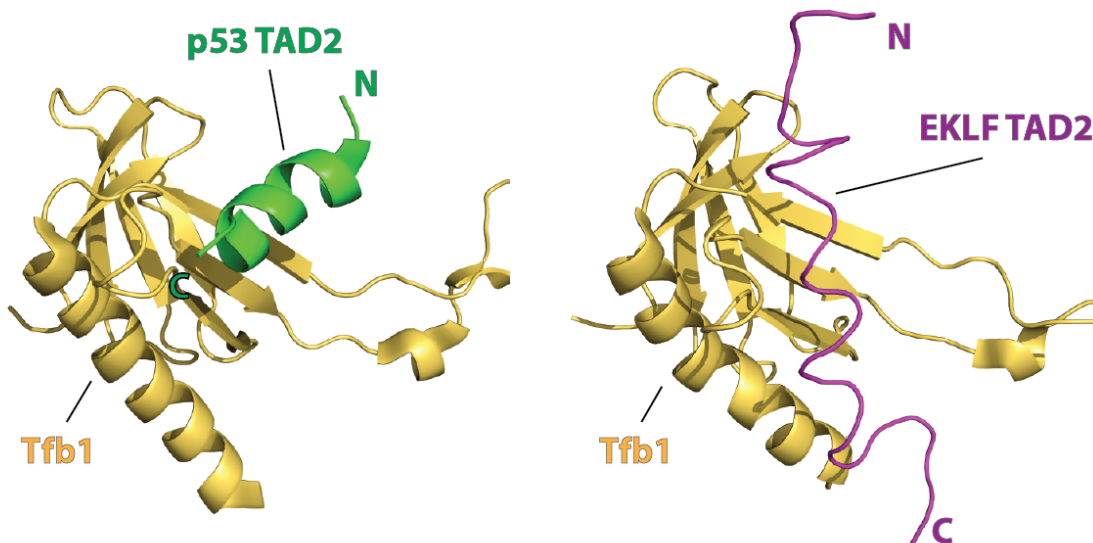
Il est intéressant de constater que le TAD de p53 est similaire à celui d'EKLf tant au point de vue de la séquence que de la fonction. Tout comme le TAD de p53, celui d'EKLf peut être divisé en deux sous-domaines qui possèdent des rôles distincts et capables d'activer indépendamment la transcription. L'alignement des séquences de p53TAD1 et p53TAD2, avec EKLfTAD1 et EKLfTAD2 respectivement, révèle qu'à chacune des trois positions hydrophobes clés des résidus de p53 correspond un résidu hydrophobe dans EKLf (**Figure 1.6 A**). De plus, plusieurs sites de phosphorylation de p53 correspondent à des sites potentiels de phosphorylation d'EKLf. Ainsi, vu ces ressemblances entre les TAD d'EKLf et ceux de p53, ceci laisserait croire que ces TAD pourraient lier les mêmes partenaires. Jusqu'à présent, le seul exemple qui appuie cette hypothèse a été démontré par le fait qu'EKLfTAD2, tout comme le p53TAD2, est capable de lier le domaine homologue de plekstrin (PH) de la protéine Tfb1. Cependant, la structure RMN des deux interactions démontre que malgré la similitude entre les deux séquences de TAD, ils n'adoptent pas la même conformation lorsqu'ils se lient. EKLfTAD2 lie le domaine PH de Tfb1 avec une conformation étendue par rapport à une conformation en hélice- $\alpha$  pour p53TAD2 (**Figure 1.6 B**) [16].



D'ordre général, les TAD acides peuvent lier un grand nombre de partenaires. Cette caractéristique est d'une part très intéressante, puisque cette faculté de changer de conformation aussi facilement peut être utile dans le design de peptides à fonctions multiples ou encore dans la conception de nanomatériaux intelligents [51, 52]. Par contre, il est difficile de prédire comment une séquence si versatile peut réguler la transcription avec une aussi bonne efficacité. Au fil des années, nous avons été capables d'analyser quelques-uns des mécanismes qui régulent les AT par les TAD acides. De tels contrôles impliquent la régulation de leur abondance, de leur localisation cellulaire et de leur activité à travers des modifications post-traductionnelles. Cependant, des études récentes ont amené à remettre en question notre compréhension du modèle courant de la régulation des TAD acides puisqu'il semblerait que les mécanismes responsables pour leur dégradation chevauchent ceux impliqués dans leur rôle d'activateur de la transcription.

**A**

|           |    |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |    |
|-----------|----|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|----|
| EKLF TAD1 | 19 | P | F | P | D | T | Q | D | D | F | L | K | W | W | R | S | E | E | A | Q | D | M | 39 |
| p53 TAD1  | 11 | E | P | P | L | S | Q | E | T | F | S | D | L | W | K | L | L | P | E | N | N | V | 31 |
| EKLF TAD2 | 64 | D | D | E | R | G | A | D | A | T | W | D | L | D | L | L | L | T | N | F | S | G | 84 |
| p53 TAD2  | 39 | A | M | D | D | L | M | L | S | P | D | D | I | E | Q | W | F | T | E | D | P | G | 59 |

**B**

**Figure 1.6. Comparaison entre les TAD de p53 et d'EKLF.**

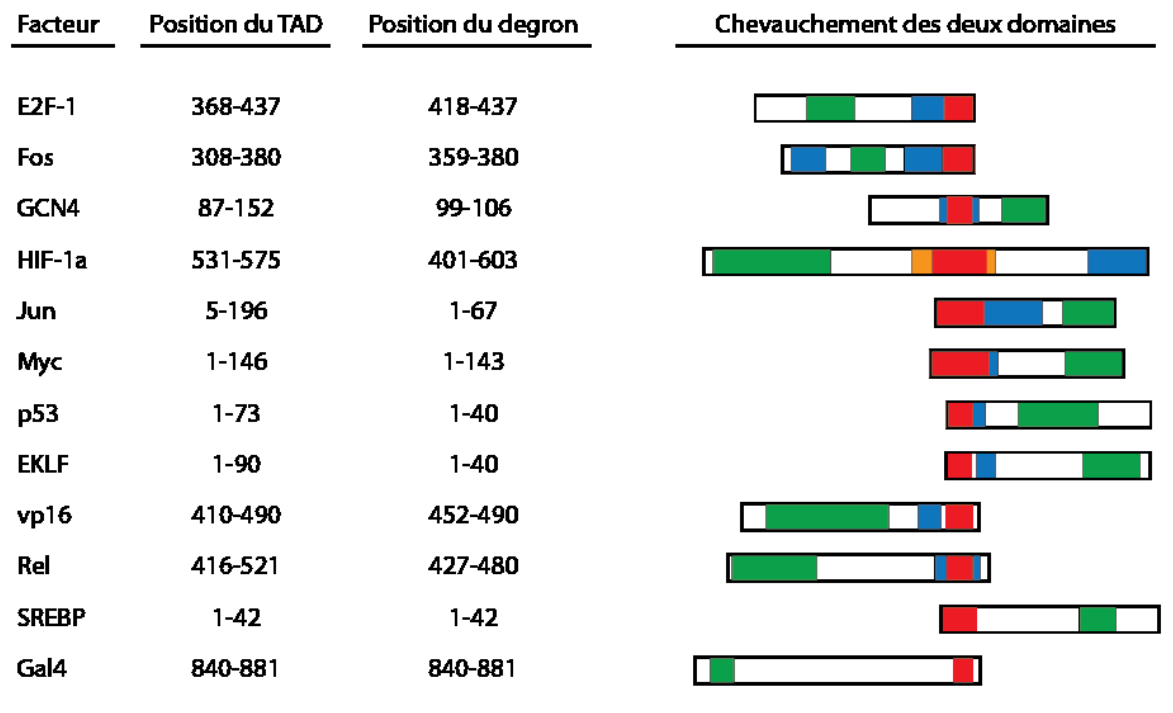
(A) Alignement des séquences TAD1 et TAD2 d'EKLF avec les TAD respectifs de p53. Les résidus connus pour former une hélice sont soulignés en noir. Les résidus hydrophobes impliqués dans les interactions sont surlignés en gris. Les sites potentiels de phosphorylation sont surlignés en jaune.

(B) Structure RMN des complexes entre le domaine PH de Tfb1 (beige) et p53 (vert) ou EKLF (mauve). Figure adaptée de [16] et réalisée avec PyMol [53] à partir des fichiers de la base de données PDB (Protein Data Bank) [54]:2GS0 et 2L2I.

### **1.1.6      *Fonction paradoxale des TAD acides***

En 2000, Tansey W. P. et son équipe ont découvert un phénomène qu'ils ont appelé « activation par la destruction » [55]. Ils ont démontré que la protéolyse d'AT possédant un TAD acide les aide à activer la transcription [55]. En d'autres mots, l'activation par le TAD acide dépend de sa dégradation. Les premiers indices de l'existence de ce phénomène paradoxal proviennent de la découverte que la région TAD acide de l'AT Gal4 chez la levure, recouvre la région signalant sa dégradation (aussi appelé « degron ») [55]. Ce recouvrement de séquence à fonction paradoxale, qu'ils ont plus tard nommé domaine d'activation/dégradation (degradation activation domains : DAD) n'est pas un cas unique à l'AT Gal4; plus de 30 AT, incluant p53 et EKLF, possèdent également ce recouvrement (**Figure 1.7**) [55-59].

Le phénomène d'activation par la destruction semble être un aspect important et récurrent dans la régulation de l'activité chez les AT possédant un TAD acide. Cependant, l'association de ces deux fonctions à l'intérieur du même domaine soulève la question : quelle est l'utilité d'avoir deux fonctions opposées aussi intimement reliées? Les premières hypothèses suggèrent que cela permettrait à la cellule de garder un contrôle sur ces activateurs en tout temps [60]. Par contre, le mécanisme par lequel ce contrôle s'effectue est encore largement méconnu. Les récents progrès dans ce domaine indiquent que la fonction paradoxale des TAD acides serait principalement contrôlée par des éléments du système de l'ubiquitine (UBI) [60].



**Figure 1.7. Les TAD et les « degon » des activateurs de transcription.**

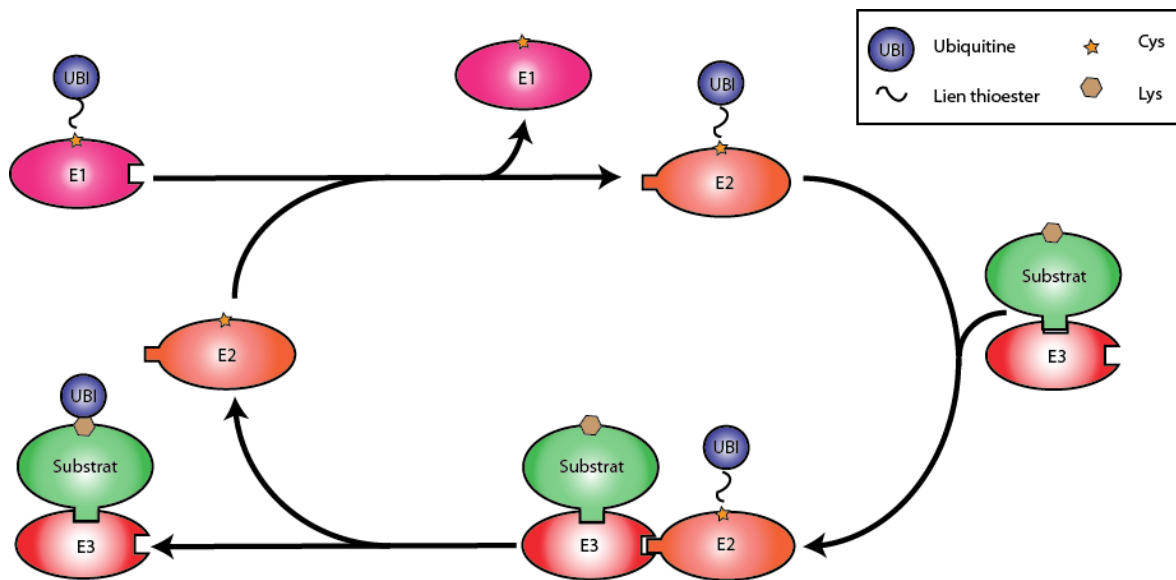
Relation entre les TAD et les régions « degon » dans différents AT. À gauche sont indiqués différents facteurs de transcription à TAD acide avec la position séquentielle de leur TAD et celle de leur « degon ». À droite sont représentés de manière schématique les différents domaines des AT avec le domaine de liaison à l'ADN en vert, le domaine TAD en bleu, le « degon » en orange, et le chevauchement entre le TAD et le « degon » en rouge. Figure adaptée de [60].

## **1.2 Le système de l'ubiquitine**

Toutes les protéines d'une cellule seront au moins une fois en contact avec le système UBI durant leur vie. Le système d'UBI a longuement été comparé au « baisé de la mort » dans le monde des protéines puisque sa liaison à des protéines cibles en tant que modification post-traductionnelle entraîne généralement la destruction de celle-ci par la machinerie protéolytique [61]. Or, malgré son omniprésence dans le contrôle de la dégradation des protéines par ubiquitination, le système UBI peut aussi jouer un rôle important à plusieurs autres niveaux tels la signalisation cellulaire, la réparation de l'ADN et la transcription à travers des interactions non-covalentes entre l'UBI et d'autres protéines [62, 63].

### ***1.2.1 L'ubiquitination***

L'ubiquitination est une modification post-traductionnelle aboutissant à la fixation covalente d'une ou plusieurs molécules d'UBI via leur queue C-terminale sur des protéines cibles. Cette fixation est un processus qui se fait de manière coordonnée avec l'aide de trois types d'enzymes dénommées enzyme activatrice (E1), enzyme de conjugaison (E2) et enzyme ligase (E3) (**Figure 1.8**) [64, 65]. Le processus débute par l'activation de l'UBI par E1. Par la suite, l'UBI activée est transférée au site actif de E2 qui fonctionne de manière synchronisée avec E3 pour conjuguer l'UBI à un groupement amino d'une protéine cible. La spécificité du processus est principalement gouvernée par l'interaction entre E3 et son site de liaison sur la protéine cible. La dernière étape de l'ubiquitination inclue la formation d'un lien isopeptidique entre le groupement carboxylé de la glycine C-terminal de l'UBI et le groupement  $\epsilon$ -amino primaire d'un résidu lysine de la protéine cible. De plus, similairement aux autres types de modifications post-traductionnelles comme la phosphorylation, la méthylation et l'acétylation, l'ubiquitination peut être réversible via l'action d'une enzyme déubiquitinase (DUB) [66]. La réaction de déconjugaison de l'UBI par les DUB est effectuée par leur activité protéase de cystéine. La famille des DUB est la plus grande parmi les enzymes du système UBI. Plusieurs bactéries et virus encodent des DUB pour exploiter le système UBI humain pour assurer leur réplication, infection et pathogénéité [67-70].

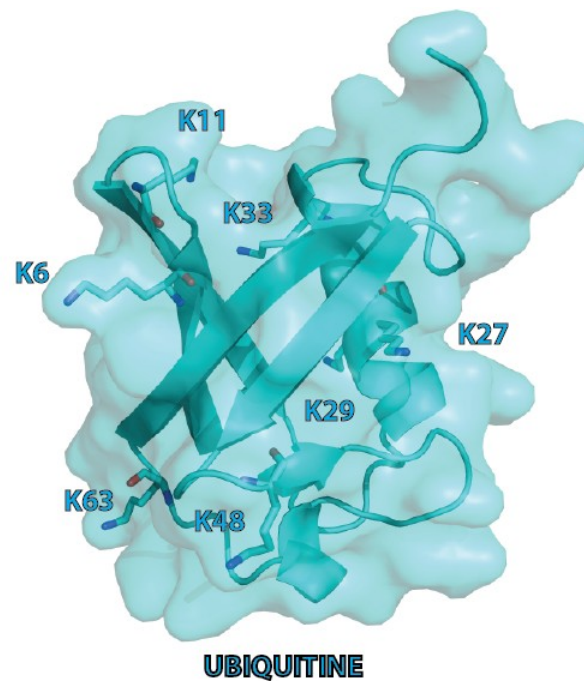


**Figure 1.8. Les différentes étapes de l'ubiquitination.**

L'ubiquitination de protéines cibles s'effectue à l'aide d'une série de réactions de transthioestérification faisant intervenir les enzymes E1, E2 et E3 du système de l'UBI. Figure adaptée de [71].

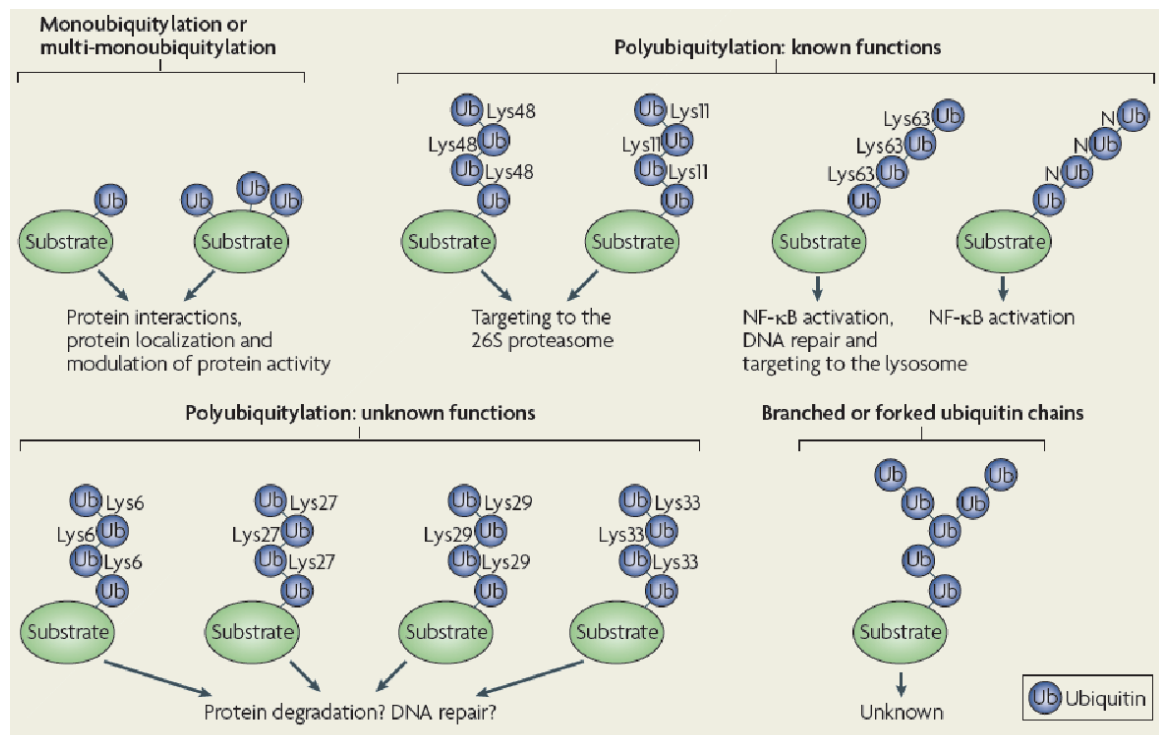
### 1.2.2 Les différents signaux de l'ubiquitine

Un aspect intéressant de l'UBI est qu'elle possède des lysines (aux positions 6, 11, 27, 29, 33, 48 et 63) qui peuvent être l'objet d'ubiquitination afin de former des chaînes poly-ubiquitines (polyUBI) (**Figure 1.9**). Ainsi, l'UBI peut s'auto-ubiquitiner pour former jusqu'à sept chaînes homotypiques et plusieurs types de constructions hétérotopiques [72]. Théoriquement, il pourrait y avoir une infinité de types de chaînes possibles, mais, en réalité, seuls quelques uns ont été identifiées dans la cellule où le type et la longueur de la chaîne formée ont un aspect déterminant pour le genre de signal transmis (**Figure 1.10**) [71]. Par exemple, une modification par une chaîne d'UBI liée en K48 cible un substrat pour sa dégradation par la machinerie protéosomale, tandis qu'une modification par les chaînes d'UBI en K63 est associée à la réparation de l'ADN et la voie de signalisation NFκB, ou encore la monoubiquitination (fixation d'une seule UBI sur la cible) est associée aux voies de signalisation impliquées dans le trafic des protéines et dans l'endocytose [73].



**Figure 1.9. Les différentes lysines présentes chez l'ubiquitine.**

Structure de l'UBI et des lysines qui peuvent être ubiquitinées. Figure réalisée avec PyMol [53] à partir du fichier PDB [74] : 1UBQ.



**Figure 1.10. Les différents signaux de l'ubiquitination.**

Schéma des différentes chaînes UBI trouvées chez l'humain ainsi que les différents processus cellulaires qui y sont associés. Figure tirée de [71].

Par contre, malgré leur lien à une fonction précise, chaque type de chaîne n'est pas limité à une seule fonction. Il a été montré que des liens sur d'autres lysines que la K48 étaient aussi capable de signaler des protéines pour leur dégradation [75]. En effet, les chaînes en K63, qui procurent normalement un signal d'activation plutôt que de dégradation, sont capables de cibler des protéines pour être dégradées [76]. Néanmoins, le fait que l'ubiquitination puisse former plusieurs types de chaînes UBI et varier leur longueur rend cette modification post-traductionnelle très flexible pour réguler plusieurs événements cellulaires.

### **1.2.3 La reconnaissance de l'ubiquitine et les domaines de liaison à l'UBI.**

La variété des événements cellulaires qui sont régulés par l'UBI suggère l'existence de protéines dans la cellule qui interagissent avec les chaînes polyUBI. En effet, il y a un nombre important de protéines capables d'interagir avec l'UBI à travers des domaines de liaison à l'UBI (Ubiquitin Bindin Domain; UBD). Il a été démontré que ces domaines formant des liaisons non-covalentes avec l'UBI jouent un rôle clé dans le contrôle de plusieurs événements cellulaires en traduisant l'information venant de la protéine ubiquitinée. À ce jour, il existe près de 20 familles d'UBD qui comptent plus de 150 protéines différentes. Les différentes classes sont les domaines hélices alpha, les domaines doigts de zinc, les domaines homologues de plekstrin (PH), les domaines « E2 conjugating-like » et les non classés [77]. Tout comme les types d'embranchement d'UBI, chaque famille d'UBD influence le destin d'une protéine.

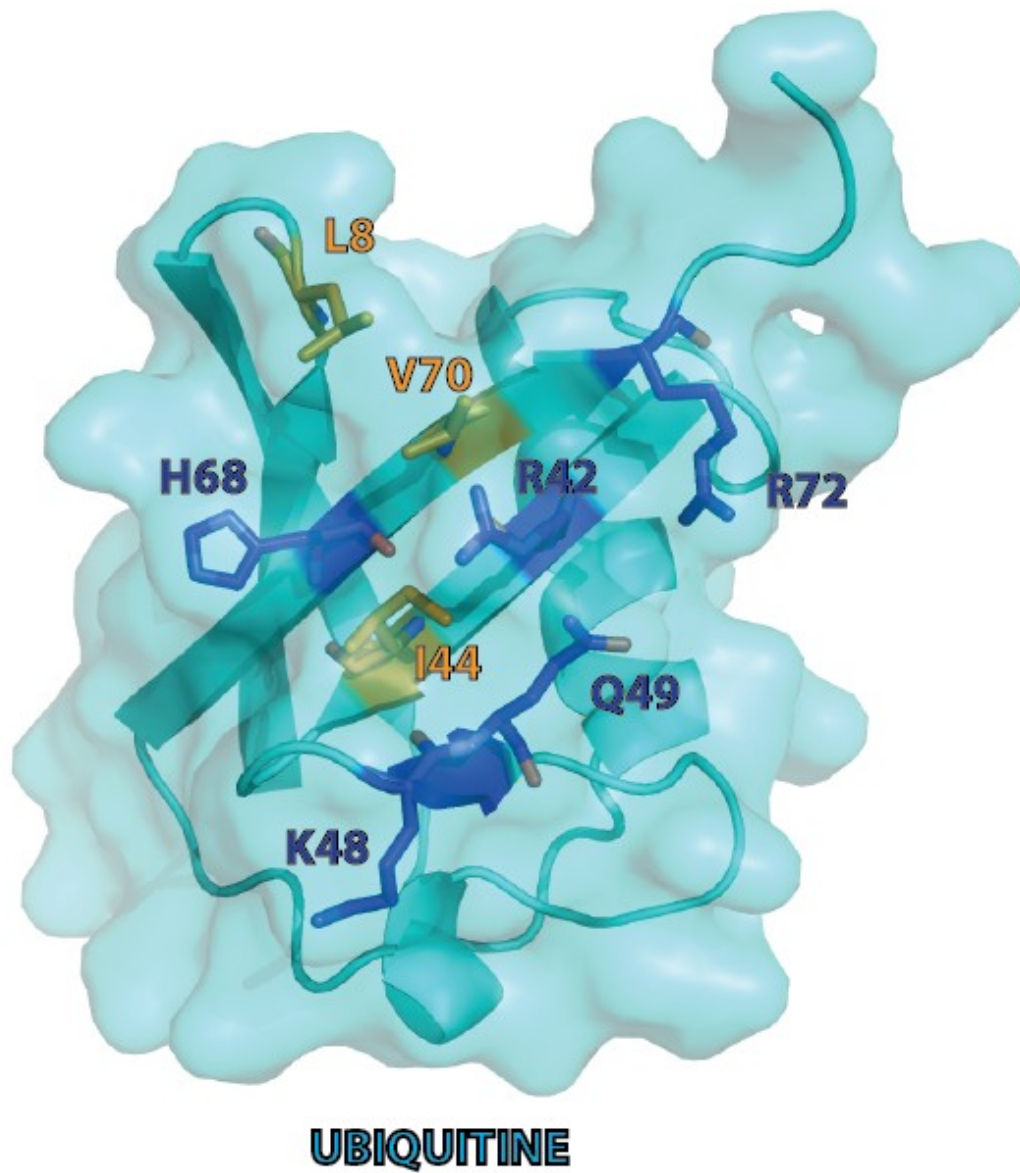
Le premier motif UBD à avoir été caractérisé était celui d'une sous-unité du protéasome : la protéine à interaction multiple avec l'UBI S5a (S5A) [78]. La séquence de S5A qui est nécessaire et suffisante pour interagir avec l'UBI est courte et simple et a été utilisée comme point de départ dans plusieurs recherches bio-informatiques afin d'identifier des séquences similaires dans d'autres protéines. Des modèles statistiques et des recherches itératives dans des bases de données basées sur la séquence de S5A ont permis d'identifier la première séquence consensus UBD connue sous le nom de motif UBI interactant (Ubiquitin Interacting Motif: UIM) [79]. Cette séquence consensus identifiée



est  $\Phi$ -x-x-Ala-x-x-x-Ser-x-x-Ac où  $\Phi$  sont des résidus hydrophobes, Ala sont des résidus alanine, Ser sont des résidus sérine, Ac sont des résidus acides et x sont n'importe quel acide aminé. Des études *in vivo* ont démontré par la suite que les protéines trouvées *in silico* contenant cette séquence sont effectivement capables d'interagir avec UBI [80-83]. Parallèlement, un deuxième motif, le domaine associé à l'UBI (Ubiquitin-Associated: UBA), a été identifié à partir de recherches bio-informatiques comme étant une région homologue dans une cohorte de protéines impliquées dans les réactions d'ubiquitination ou de dé-ubiquitination [84-86]. Ces domaines sont caractérisés par la formation de triple hélices qui interagissent avec l'UBI au niveau de leur séquence hydrophobe conservée présente sur la première hélice [84]. La découverte des domaines UIM et UBA a provoqué par la suite un ouragan d'études sur les UBD. Parmi les UBD identifiés, tous sont relativement petits (15-150 acides aminés), capables d'interagir avec la monoubiquitine, mais ont généralement une préférence pour un type de chaîne polyUBI avec des substrats ubiquitinés [87].

#### **1.2.4 La surface d'interaction L8-I44-V70 de l'ubiquitine**

La plupart des UBD interagissent au niveau de la surface hydrophobe de l'UBI. Des études de mutagenèse portant sur l'interaction UBI-UBD révèlent que les résidus les plus importants au sein de la surface de contact de l'UBI sont les trois résidus hydrophobes L8, I44 et V70 [77, 88, 89]. L'interaction est parfois facilitée par l'habileté des résidus hydrophiles dynamiques avoisinant cette surface hydrophobe (R42, K48, Q49, H68 et R72) (**Figure 1.11**) [62, 90-92]. De ce fait, si les UBD se lient toujours au même endroit sur l'UBI, comment font-ils pour distinguer les différentes chaînes d'UBI? Bien que les UBD lient la surface hydrophobe de l'UBI, les résidus spécifiques au cœur de l'interaction peuvent varier créant une empreinte d'interaction spécifique pour chaque UBD. Ces empreintes d'interaction permettent aux UBD d'avoir une préférence pour les différentes chaînes d'UBI [86, 93, 94].



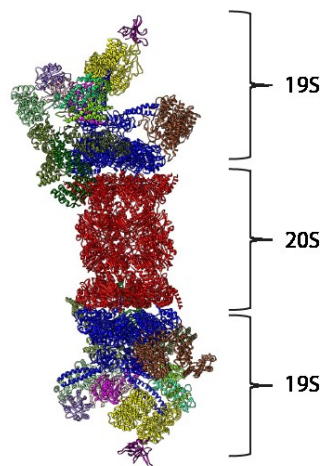
**Figure 1.11. La surface d'interaction L8-I44-V70 de l'ubiquitine.**

Schéma de l'UBI et des résidus impliqués dans la majorité des interactions avec des UBD. Les résidus hydrophobes L8, I44 et V70 (bâtonnets orange) sont les plus importants tandis que les résidus hydrophiles R42, K48, Q49, H68 et R72 (bâtonnets bleus) servent à faciliter l'interaction. Figure réalisée avec PyMol [53] à partir du fichier PDB [74] : 1UBQ.

### 1.3 Le système d'ubiquitine dans la régulation des activateurs

#### 1.3.1 *Rôle protéolytique*

Dans la transcription, l'ubiquitination d'activateurs sert à contrôler l'activité de ces derniers de plusieurs façons. L'une de ces façons est le contrôle protéolytique où les niveaux d'activateurs sont régulés par la coopération entre l'ubiquitination des AT et leur dégradation par le protéasome [95]. Cette coopération, mieux connu sous le nom de système ubiquitine/protéasome (ubiquitin/proteasome system: UPS) fonctionne de telle sorte que les protéines qui sont ubiquitinées par une chaîne polyUBI spécifique, seront reconnues puis dégradées par le protéasome. Chez les eucaryotes, le protéasome comporte deux sous-unités à fonction bien distinctes [96]. La sous-unité cylindrique 20S située au cœur de la fonction du protéasome qui s'occupe du passage des substrats ubiquitinés et de leur dégradation dans le protéasome à l'aide de plusieurs protéases, et les sous-unités 19S de chaque côté du protéasome qui forment une base d'ATPases capables d'activer la transcription (**Figure 1.12**) [97, 98]. La sous-unité 19 S possède aussi plusieurs enzymes DUB et des E3 ligases, suggérant que le débit de dégradation de protéines par le protéasome est contrôlé par un équilibre dynamique entre l'ubiquitination et la déubiquitination [99].



**Figure 1.12. Le protéasome 26S.**

Représentation de la structure du protéasome 26S et de ces sous-unités 19S et 20S. Figure adaptée de [100].

Outre le fait que l'UPS soit capable de contrôler l'abondance des AT dans la cellule, certains exemples démontrent qu'il est aussi capable de contrôler la localisation des activateurs ou la régulation de leur association à des co-activateurs [101]. D'autres expériences, chez la levure, montrent que l'UPS peut réguler la transcription en contrôlant le temps qu'un AT est actif. En effet, certains AT requièrent l'ubiquitination pour bien fonctionner malgré le fait que ce même signal les cible pour être dégradés par le protéasome [102]. Ceci suggère que l'ubiquitination agit comme bombe à retardement pour assurer que l'AT n'active pas la transcription éternellement.

### **1.3.2 Rôle non-protéolytique**

Le système de l'UBI peut aussi contrôler les activateurs à travers un mécanisme non-protéolytiques qui évite l'UPS, entre autres par la monoubiquitination. En effet, l'attachement d'une seule UBI à un activateur est capable d'influencer sa fonction indépendamment de sa protéolyse [103-106]. Des études ont prouvé que la monoubiquitination peut, dans certains cas, aider un activateur à s'ancrer sur la chromatine, alors qu'en d'autres cas, elle les aiderait à s'en défaire [107, 108]. En effet, la monoubiquitination de l'AT Gal4 empêche son relâchement de la chromatine ce qui lui permet d'exercer son activité plus longtemps [107]. Par contre, chez la levure, la monoubiquitination des activateurs de la famille  $\alpha 2$  induit le recrutement de la chaperonne ATPase Cdc48 qui extrait ces activateurs de la chromatine pour les inactiver [108-110]. D'autres études démontrent que la monoubiquitination d'un activateur peut être utile même avant son association à la chromatine où l'ajout de monoUBI peut avoir des effets inverses sur différents AT [102, 105]. Outre la monoubiquitination, d'autres types de chaînes polyUBI seraient également importantes dans le contrôle des activateurs comme les polyUBI K63 ainsi que les polyUBI atypiques qui sont beaucoup moins connus [111]. Ainsi, ces exemples nous permettent de constater qu'il n'y a pas de mécanisme de contrôle commun pour tous les activateurs.

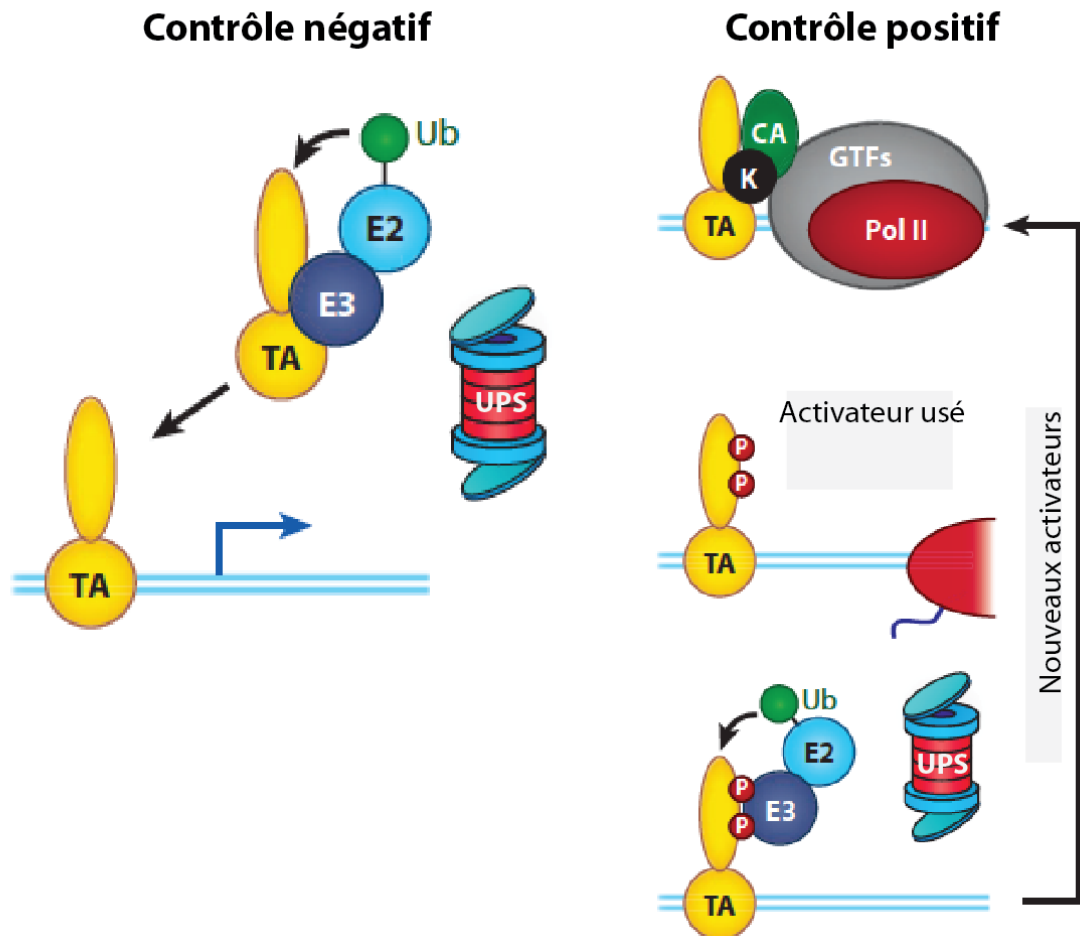
### **1.3.3      *Le système de l'ubiquitine et les deux modèles du mécanisme de la fonction paradoxale des activateurs***

Il n'y a pas encore de consensus pour expliquer le mécanisme de la fonction des domaines paradoxaux des TAD acides. En effet, il est difficile de déterminer le fonctionnement de ces domaines puisque certains fonctionneraient différemment. Cependant, étant donné que leur fonction soit dépendante de la dégradation de l'AT, plusieurs pensent que les mécanismes de contrôle des TAD acides doivent connecter avec l'UPS. Ainsi, deux modèles, qui réfèrent à un concept « kamikaze » d'AT où la destruction de l'activateur est une étape obligatoire du mécanisme d'activation, ont été proposés: (1) le modèle d'activation négative et (2) le modèle d'activation positive. (**Figure 1.13**) [112-115].

Le modèle du contrôle d'activation négative stipule que l'UPS empêche la transcription de se produire en éliminant l'AT avant sa liaison à la chromatine. Dans ce modèle, l'UPS limite l'activité des AT en régulant leur abondance par la liaison directe d'une E3 ligase sur la section DAD qui initie sa protéolyse via l'ubiquitination. Si la protéolyse ne prend pas place, les AT s'accumulent ce qui résulte en une augmentation de la transcription. Ce type de contrôle s'appliquerait, entre autres, à l'activateur p53 par sa liaison avec la E3 ligase MDM2. D'autres activateurs comme la protéine virale 16 (VP16) du virus de l'herpès et la protéine Gal4 de la levure fonctionnent aussi de cette manière.

Le deuxième modèle, celui du contrôle d'activation positive, stipule que l'UPS aurait un effet bénéfique sur la transcription en éliminant des AT usés qui bloqueraient les sites de transcription. Principalement basé sur des résultats obtenus sur l'AT « Sterol regulatory element-binding protein 1a » (SREBP1a) [116] qui est requis pour la biosynthèse de cholestérol et d'acide gras, ce modèle propose que des kinases présentes dans la machinerie transcriptionnelle serviraient à marquer les activateurs comme étant usés lorsque leur fonction est terminée. Cette phosphorylation servirait en même temps à recruter une E3 ligase qui enclencherait le processus de protéolyse par l'UPS de l'AT non-fonctionnel qui bloque le site de la transcription. Cependant, ce modèle reste incomplet puisqu'il ne permet pas d'expliquer plusieurs données expérimentales. Notamment, sachant que la liaison de la E3 ligase sur l'AT ne se fait pas au niveau du DAD dans ce type

de régulation, le modèle de régulation positive n'explique pas pourquoi la délétion du domaine DAD de SREBP1a diminue grandement sa dégradation. D'autres DAD dans cette catégorie, comme le DAD d'EKLf, voient aussi leur dégradation diminuée lorsque leur DAD est enlevé malgré la persistance de leur ubiquitination. Ainsi, ceci indiquerait qu'il manquerait des étapes de régulation dans le modèle de contrôle positif.



**Figure 1.13. Les deux modèles du mécanisme de contrôle des domaines paradoxaux.**

Représentation des deux modèles du mécanisme de contrôle des AT. Le modèle de contrôle négatif suggère que l'AT se fait réguler par l'UPS avant sa liaison à la chromatine en contrôlant son niveau. Le modèle de contrôle positif suggère que des kinases de la machinerie transcriptionnelle de base phosphorylent l'AT pour l'inactiver et, du même coup, bloquer la transcription. Cette phosphorylation servirait aussi à recruter des E3 ligases qui mèneraient à la dégradation de l'AT afin de libérer le site pour une transcription subséquente. Abréviations : CA, coactivateur; E2, enzyme de conjugaison de l'UBI; E3, enzyme de ligase UBI; TA, activateur de transcription; UB, ubiquitine; P, phosphorylation; K, kinase; GTF, facteurs généraux de transcription; lignes bleues, l'ADN. Figure tirée de [5].

#### 1.4 Hypothèse et objectifs

Le phénomène d'activation par la destruction entrepris par les DAD est un aspect important et récurrent dans la régulation de l'activité chez une grande partie des AT possédant un TAD acide. Depuis la découverte du chevauchement des séquences de dégradation et d'activation sur les TAD acides à partir de l'AT chez la levure Gal4, des progrès significatifs ont été faits pour trouver plusieurs AT importants chez l'humain qui font partie de cette catégorie, comme p53, SREBP1a et EKLF, mais aussi chez les virus comme la protéine virale 16 (VP16) du virus de l'herpès. Bien qu'il y ait des modèles pour la régulation de ces DAD, le concept de l'activation par la destruction est encore relativement nouveau. De ce fait, les mécanismes précis de la régulation des DAD restent encore peu connus, et ce, surtout au niveau des DAD faisant partie du modèle de régulation positive.

Un aspect intéressant des DAD à l'intérieur des TAD acides est que leur forte teneur en résidus hydrophobes et acides est similaire à celle de certaines familles d'UBD. Plus précisément, les DAD ressemblent aux UBD de la famille UIM ainsi que leurs variants motifs interagissant avec l'UBI (motifs interacting with ubiquitin: MIU) [79]. En effet, les éléments clés de la reconnaissance de ces domaines UBD à l'UBI partagent le motif  $\Phi \Phi X X \Phi$  (où  $\Phi$  est un résidu hydrophobe et X n'importe quel résidu) retrouvé dans plusieurs domaines DAD (**Figure 1.14**).

**UIM**

|         |           |               |
|---------|-----------|---------------|
| S5a     | (222-210) | SVRLALALEPDAS |
| Hrs     | (272-260) | SQSLALALQLEEE |
| Vps27-1 | (272-260) | KLSEIAKRILEE  |
| Vps27-2 | (315-299) | RLSEQIAAKLDPD |
| STAM-1  | (185-169) | KLSEIAKALDEE  |
| STAM-2  | (179-163) | QLSEIAKAIDEE  |

**MIU**

|           |             |               |
|-----------|-------------|---------------|
| RNF168-1  | (173-187)   | KSDEELARKLSID |
| RNF168-2  | (444-458)   | EQDRLLALQLQKE |
| Myosin VI | (1003-1017) | EAHQQALDDLQVE |
| Rabex5    | (68-56)     | QEDWELAERLQRE |

**DAD**

|         |           |               |
|---------|-----------|---------------|
| EKLF    | (22-34)   | DTQDDFLKWRSE  |
| p53     | (17-29)   | ETFSDLWKLLPEN |
| SREBP1a | (27-39)   | TDIEDMLQLINNQ |
| Ga14    | (859-871) | TTMDDVYNLYFDD |
| VP16    | (431-445) | MTPADALDDFDLE |

-----ΦΦXXΦ-----

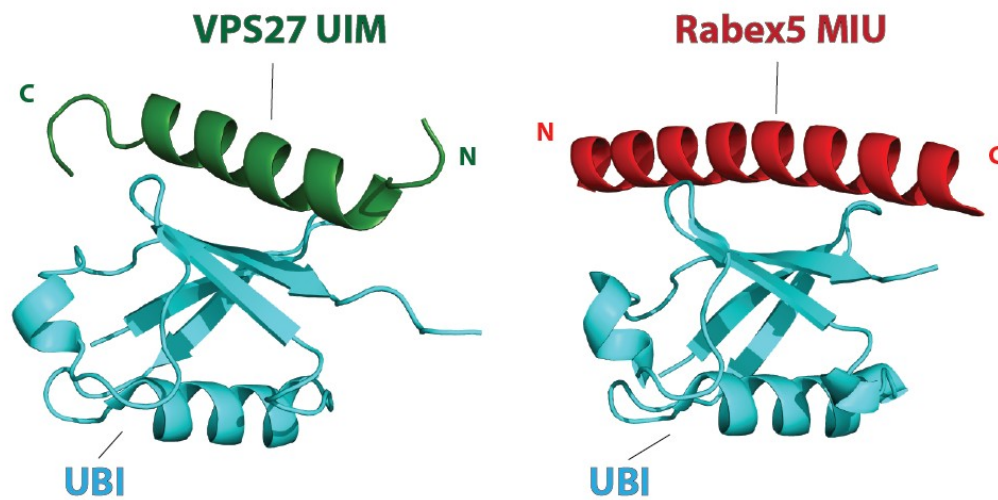
**Figure 1.14. Comparaison entre les séquences d'UIM, MIU et DAD.**

Les résidus hydrophobes et acides sont colorés en bleu et rouge respectivement. Φ représente un résidu hydrophobe et X un résidu quelconque.

Les UIM et les MIU sont des domaines principalement trouvés dans des protéines de transport qui reconnaissent des protéines ubiquitinées. Par exemple, les UIM, bien que retrouvés dans le protéasome par la présence de S5a (section 1.2.3), ils sont aussi dans des protéines du transport des vésicules d'endocytose comme Hrs, Stam et VPS27. Les MIU, eux, sont aussi retrouvés chez des protéines qui régulent le transport des vésicules d'endocytose comme dans Rabex5 ou Myosin VI, mais peuvent aussi être retrouvé dans des protéines E3 ligases comme le démontre les deux domaines MIU de RNF168.



Cependant, il est à noter que, parmi les protéines possédant des MIU seul celui de Rabex 5 a été caractérisé structuellement. La différence principale entre les UIM et les MIU est le fait que le domaine MIU lie l'UBI avec une conformation inversée par rapport à la liaison de l'UBI à UIM malgré que les deux domaines lient la même surface hydrophobe de l'UBI (**figure 1.15**).



**Figure 1.15. Comparaison entre les domaines UIM et MIU.**

Structures des complexes entre UBI (en cyan) et le domaine UIM de VPS27 (en vert) ou le domaine MIU de Rabex5 (en rouge). Figure réalisée avec PyMol [53] à partir des fichiers PDB [74] : 2FIF et 1QOW.

Les UIM et MIU sont souvent présents en tandem sur la protéine et ils sont, dans plusieurs cas, séparés par une séquence hautement dynamique. Par exemple, Vps27 est une protéine qui contient deux domaines UIM (Vps27-1 et Vps-2) séparés par un espaceur dynamique sur sa séquence, ce qui lui permet d'interagir avec plusieurs formes de signaux UBI [117]. Dans leur état libre, certains UIM et MIU adoptent une structure secondaire en hélice  $\alpha$  qui est à la base de la reconnaissance de l'UBI où ils lient les résidus L8-I44-V70 de la surface hydrophobe de l'UBI (**Figure 1.15**). Or, malgré la présence des hélices chez les UIM et MIU, ils ne sont pas organisés en une structure tertiaire à l'état libre. La présence de plusieurs régions sans structure tertiaire à l'état libre est similaire à ce qu'on peut voir sur des domaines TAD acides et leur permet donc de lier plusieurs types de signaux UBI permettant de jouer différents rôles au sein de la cellule. Ainsi, ces fortes similarités entre

les UIM, MIU et les DAD amènent à se poser la question : **est-ce que les DAD pourraient agir comme un domaine UBD?**

Étant donné la prévalence des DAD chez les AT à TAD acides et leurs impacts sur le contrôle de la transcription, il est essentiel de connaître plus en détail le fonctionnement de ce domaine. C'est dans le but d'améliorer notre compréhension de la régulation de la fonction des TAD acides par leur DAD que nous avons entrepris des études structurales et fonctionnelles de plusieurs interactions non-covalentes impliquant ces domaines et l'UBI. Dans ce contexte, ce travail porte sur l'étude de la reconnaissance de l'UBI par les domaines TAD acides d'AT et contient trois objectifs:

### **1) Déterminer si les DAD peuvent lier l'UBI.**

Étant donné la forte similarité entre les DAD et les UIM ou MIU, il est certainement possible que le domaine paradoxal des TAD acides agisse comme un UBD pour réguler sa fonction et que ceci pourrait venir ajouter une pièce manquante aux modèles de régulation des DAD.

Dans cette étude, j'ai utilisé la Résonance Magnétique Nucléaire (RMN) pour déterminer si les DAD de TAD acides des deux types de modèle de régulation sont capables de former une liaison non-covalente avec l'UBI. Les DAD du modèle positif de régulation testés sont: ceux du gardien du génome p53, de VP16 qui joue un rôle clé au niveau de la transition du stade latent au stade actif du virus de l'herpès [118] et de l'AT chez la levure sur lequel les DAD sont basés Gal4. Les DAD du modèle négatif de régulation testés sont: ceux de SREBP1a qui est à la base de ce modèle de régulation et aussi du facteur de différenciation hématopoïétique EKLF.

### **2) Caractériser la structure formé par un de ces complexes.**

Dans le cas où le premier objectif démontre l'existence d'une interaction entre un DAD et l'UBI, la détermination de la structure d'un des complexes formés

permettra de déterminer quels sont les résidus importants pour l'interaction et aussi de comparer la structure avec celle des domaines UIM et MIU pour qui ils partagent beaucoup de ressemblance. La RMN, la technique la plus utilisée pour la caractérisation structurale de domaines fonctionnels très flexibles comme les TAD acides, a été choisie pour l'étude de complexes entre les DAD et l'UBI. Par contre, étant donné le temps que ça prenne pour résoudre la structure d'un complexe par RMN, advenant qu'ils lient l'UBI, on s'en est tenu à résoudre la structure d'un DAD modèle par mécanismes de régulation qui sont: p53 pour le mécanisme de régulation positive et EKLF pour le mécanisme de régulation négative. C'est deux modèles ont été préférés par rapport aux autres DAD testés puisque nous avons travaillé souvent avec ces deux protéines dans le passé [8, 10, 16, 119, 120].

### **3) Caractériser fonctionnellement le complexe entre les DAD et l'UBI en se basant sur son importance dans la stabilité de la protéine possédant le DAD.**

Les domaines TAD acides qui contiennent les DAD ont été amplement étudiés dans le passé pour leur fonction d'activation de la transcription. Cependant, ceci n'est pas le cas pour leur fonction de dégradation. En effet, bien que quelques données existent sur l'effet d'un DAD sur la stabilité de l'AT qui le contient, la grande partie de ces études portent sur l'effet d'une délétion complète du domaine ne permettant pas d'identifier spécifiquement les résidus clés comme, par exemple, ceux du motif  $\Phi \Phi X X \Phi$  des DAD. Ainsi, pour corroborer la fonction d'un DAD avec les résultats obtenus par RMN au deuxième objectif, des essais de dégradation se basant sur la luciférase serviront à mesurer la stabilité de l'AT possédant un DAD. Les résidus sur le DAD qui vont être observés par RMN comme étant impliqués dans l'interaction seront mutés afin de déterminer leurs effets sur l'interaction UBI/DAD et sur la stabilité des AT qui possèdent ces DAD.

## **1.5 Techniques utilisées**

La technique la plus commune pour étudier la structure de protéines seules ou en interaction avec un partenaire est la cristallographie aux rayons X. Sur les 96 000 structures qui sont accessibles dans la PDB, 88 % ont été résolues par cristallographie [54]. Cette technique permet de résoudre la structure à l'échelle atomique de molécules de toutes tailles, à condition qu'elles cristallisent; ce qui représente le défi majeur de cette technique. Malheureusement, il est très difficile de cristalliser des protéines intrinsèquement désordonnées comme les TAD à cause de leur grande flexibilité. Un moyen pour contourner ce problème serait de les connecter à une séquence connue pour sa faculté à bien cristalliser, ou encore de les connecter à une molécule qui les stabiliserait. Cependant, ce genre de modification peut engendrer une panoplie de contraintes artéfactuelles de cristallisation ce qui pourrait altérer la structure ou la fonction de la protéine étudiée. Une technique alternative pour obtenir la structure d'un complexe protéique est d'utiliser la diffusion de rayons X aux petits angles (SAX). Cette technique s'effectue sur des macromolécules en solution; ce qui permet de s'affranchir de la production de cristaux, mais la résolution de la structure obtenue (1-2nm) ne permet d'obtenir que la forme générale de la protéine ce qui empêche d'effectuer des études structurales poussées.

La Résonance Magnétique Nucléaire (RMN) est l'outil idéal pour étudier les protéines ou domaines intrinsèquement désordonnés, comme les TAD acides, puisqu'elle permet d'obtenir une structure à haute résolution des protéines sans avoir à former des cristaux. La RMN est aussi un outil particulièrement sensible pour l'étude d'interactions entre protéines puisqu'elle permet de détecter une large gamme d'affinité d'interaction entre les partenaires (du  $\mu\text{M}$  au  $\text{nM}$ ), et de localiser directement l'interaction sur la protéine cible. La prochaine section servira à décrire les méthodes utilisées pour la production des protéines, pour l'étude par RMN ainsi que des essais fonctionnels pour l'étude de TAD acides et l'UBI.

### 1.5.1 Production et purification de protéines pour la RMN

Pour être capable d'étudier des interactions entre protéines en RMN, il est nécessaire d'obtenir une assez grande quantité des protéines étudiées pures non-marquées et marquées en  $^{15}\text{N}$  ou en  $^{13}\text{C}/^{15}\text{N}$ . La production de protéines peut se faire par synthèse chimique où chaque acide aminé est ajouté un par un par une série de réactions chimiques *in vitro* [121] où le rendement de la synthèse chimique est généralement indirectement proportionnel à la taille du peptide synthétisé, et ce, à cause de l'accumulation du pourcentage de rendement non complet à chaque étape. La synthèse chimique est donc considérée comme simple et efficace uniquement pour la production de peptides de moins de 100 résidus. Cependant, le coût de production de peptide marqué est très dispendieux comparativement à un échantillon non-marqué. Une alternative moins coûteuse consiste à synthétiser les protéines marquées/non-marquées à partir d'un vecteur bactérien. En transfectant un plasmide codant la protéine d'intérêt à des bactéries, il est possible de les cultiver en milieu nutritif standard ou minimal  $^{13}\text{C}/^{15}\text{N}$  pour qu'elles expriment cette protéine en grande quantité de façon marquée ou non.

*Escherichia coli* (*E. coli*) est la bactérie la plus connue et la plus utilisée pour la production de protéines recombinantes à l'aide de procédés bactériens [122]. Plusieurs lignées d'*E. coli* ont été synthétisées afin de contourner plusieurs problèmes de production comme l'insertion d'ARN de transfert pour la synthèse de codon rare, la mutation de gènes pour inhiber la dégradation d'ARNm, la délétion de gènes pour empêcher la protéolyse des protéines produites, etc. Toutefois, le système bactérien *E. coli* possède certains inconvénients, dont la formation possible de corps d'inclusion et la production élevée d'acétate qui tend à inhiber la croissance de bactéries. Pour remédier à ces problèmes, il est possible de fusionner le peptide d'intérêt à une protéine recombinante qui améliore la production. Ces protéines recombinantes confèrent aux peptides stabilité et solubilité en plus de permettre une production efficace et peu coûteuse de peptides dans un système d'expression bactérien. De plus, ces protéines peuvent être utilisées pour faciliter le processus de purification du peptide. Dans le cas spécifique des TAD, nous n'avons été capables de les produire que lorsque fusionnés à la

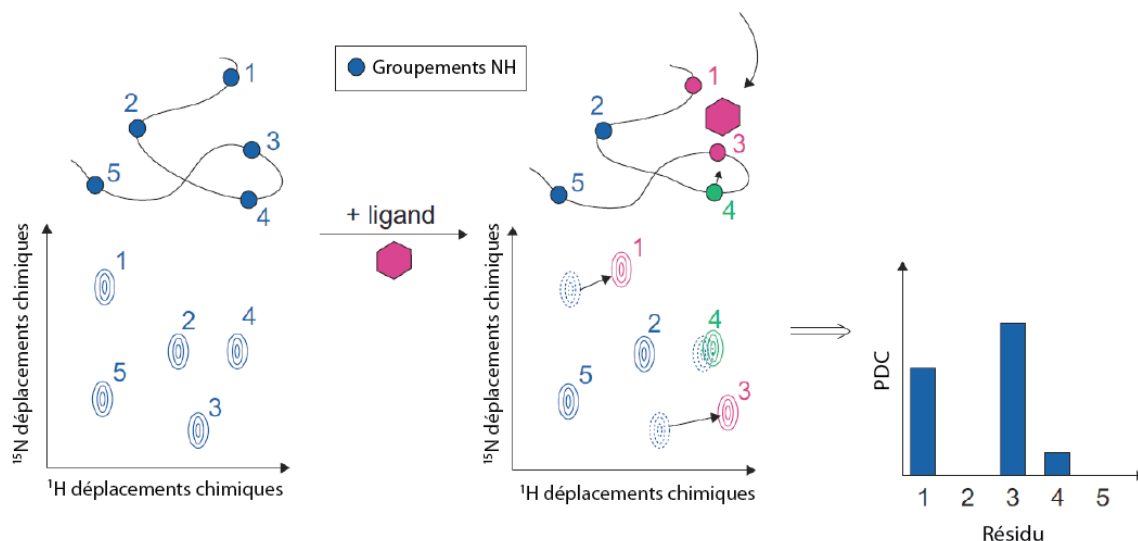
protéine Glutathione S-Transferase (GST) en N-terminal. Sans la GST, aucune des lignées bactériennes testées n'était capable de produire la protéine ou le peptide d'intérêt.

Les méthodes de purification utilisées pour purifier nos peptides sont relativement standards. D'ordre général, le protocole de purification est basé sur les propriétés intrinsèques de la protéine; point isoélectrique (PI), taille, stabilité et polarité. Ainsi, avec un mélange de méthodes de purification par chromatographie d'affinité avec la GST, d'échange d'ions, de filtration sur gel et par chromatographie en phase liquide à haute performance (CLHP), il est possible d'obtenir une grande quantité de protéines avec une très grande pureté. La pureté de chaque protéine a été vérifiée à l'aide de spectrométrie de masse.

### **1.5.2 Le titrage RMN pour détecter et caractériser les interactions entre protéines.**

Le titrage en RMN aussi appelé analyse de perturbation des déplacements chimique (PDC) en RMN est un moyen simple et efficace pour déterminer la présence d'une interaction entre protéines à l'aide de l'expérience 2D  $^1\text{H}$ - $^{15}\text{N}$  HSQC (Heteronuclear Single Quantum Coherence). Le spectre 2D  $^1\text{H}$ - $^{15}\text{N}$  HSQC permet d'observer des signaux de déplacements chimiques d'une protéine marquée  $^{15}\text{N}$  corrélant la fréquence entre une paire d'atomes liée  $^{15}\text{N}$ - $^1\text{H}$ . Ainsi, au moins un signal peut être observé pour tous les acides aminés d'une protéine marquée  $^{15}\text{N}$  (à l'exception de la proline) puisqu'ils possèdent un lien N-H sur leur chaîne principale. Les acides aminés tels que l'arginine, l'asparagine, la glutamine et le tryptophane possèdent un lien  $^{15}\text{N}$ - $^1\text{H}$  ou  $^{15}\text{N}$ - $^1\text{H}_2$  sur leur chaîne latérale qui sont également visibles et facilement identifiables à cause de la localisation de leurs signaux spécifiques. Un aspect important en RMN est que les déplacements chimiques d'un atome ou d'un lien sont particulièrement sensibles à leur environnement et donc peuvent être perturbés par de petits changements de pH, de température ou par la liaison d'un partenaire (**Figure 1.16**) [123-125]. Ainsi, en effectuant la collecte de plusieurs spectres 2D  $^1\text{H}$ - $^{15}\text{N}$  HSQC d'une protéine marquée à laquelle on ajoute une quantité croissante de protéine non-marquée, il est possible de déterminer si ces dernières

forment une interaction. Cette expérience de titrage RMN permet d'observer des PDC dans le spectre dans le cas où deux protéines sont des partenaires d'interaction. D'ordre général, les signaux qui subissent les plus gros déplacements lors du titrage sont ceux qui sont le plus impliqués dans l'interaction. Ainsi, si l'on connaît quel signal correspond à quel résidu, nous pouvons déterminer quels sont les résidus impliqués dans l'interaction. Ceci permet donc de déterminer s'il existe une interaction entre deux protéines et aussi de cartographier les résidus impliqués dans l'interaction et peut être utilisé comme contrainte structurale pour simuler la structure du complexe.



**Figure 1.16. Expérience de titrage par RMN-2D.**

Détection de la liaison d'un ligand sur une protéine par RMN-2D et par PDC dans le cas d'échange rapide. L'introduction d'un ligand induit la perturbation de l'environnement chimique de quelques résidus (résidus 1 et 3 en rose). Le ligand peut aussi introduire des modifications dans la conformation de la protéine ce qui perturbe certains résidus même s'ils ne sont pas directement au ligand (résidu 4 en rouge). L'information des PDC peut ensuite être traduite en histogramme en fonction de la séquence de la protéine pour faciliter la visualisation des résultats (à droite). Figure tiré de [125].

### 1.5.3 Les PDC pour quantifier la force d'interaction.

La force d'interaction entre deux molécules peut être vue comme la balance entre l'association et la dissociation des deux molécules à l'état d'équilibre. Cette force d'interaction est représentée par la constante de dissociation  $K_D$ . La mesure du  $K_D$  d'une interaction, lorsque jumelée à des expériences de mutagenèse, permet de mettre en valeur le rôle joué par des résidus dans l'interface de l'interaction. Elle permet aussi de

comparer la force de liaison de molécules pour une cible pour, par exemple, identifier le meilleur inhibiteur d'une cible thérapeutique.

En RMN, les  $K_D$  peuvent être déterminés en utilisant l'information des PDC obtenus lors d'une expérience de titrage RMN [126]. Dans le cas d'une interaction de stœchiométrie 1:1 entre deux protéines, l'équilibre thermodynamique ou  $K_d$  peut être décrit comme suit :

$$K_D = \frac{[P][L]}{[PL]} = \frac{k_{off}}{k_{on}} \quad [I]$$

où  $[P]$  et  $[L]$  sont respectivement la concentration de protéine et de ligand et  $[PL]$  la concentration du complexe. Dans une expérience de titrage RMN, où plusieurs spectres sont enregistrés avec des concentrations croissantes de ligand,  $[P]$  et  $[L]$  évoluent de l'état libre à l'état complexé :

$$\begin{aligned} [P]_0 &= [P] + [PL] \\ [L]_0 &= [L] + [PL] \end{aligned} \quad [II]$$

où  $[P]_0$  et  $[L]_0$  sont respectivement les concentrations de la protéine et du ligand au départ. Ainsi, en combinant les équations [I] et [II], on obtient la fonction non linéaire suivante :

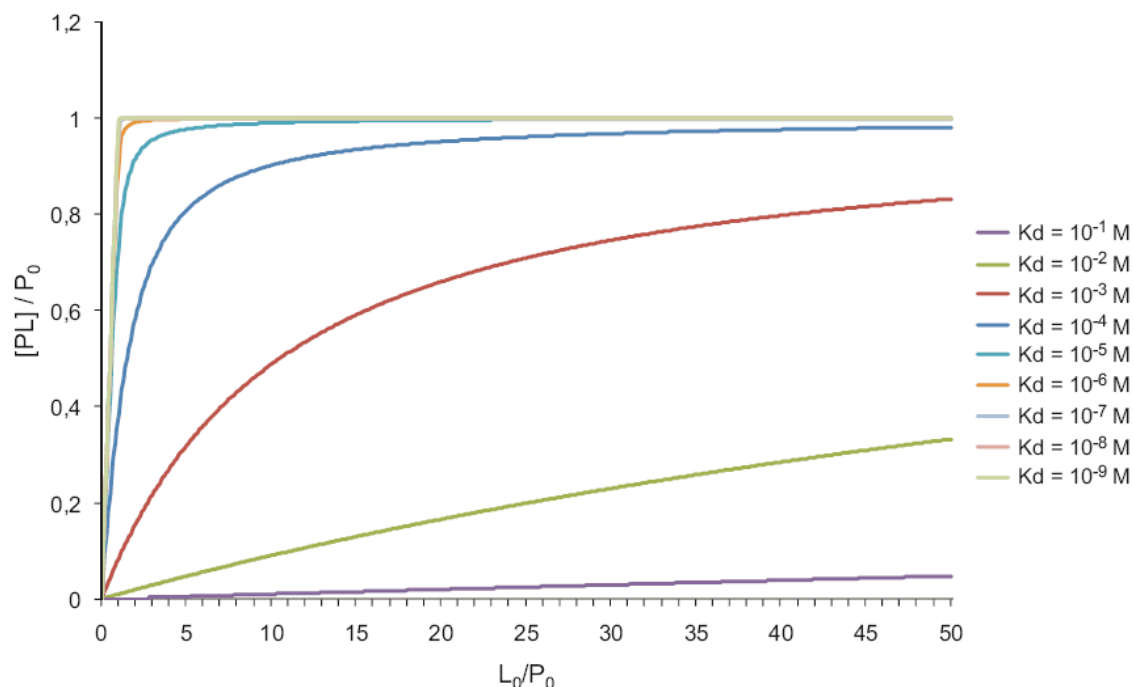
$$\frac{[PL]}{[P]_0} = \frac{1}{2} \left[ 1 + \frac{[L]_0}{[P]_0} + \frac{K_D}{[P]_0} - \sqrt{\left( 1 + \frac{[L]_0}{[P]_0} + \frac{K_D}{[P]_0} \right)^2 - 4 \frac{[L]_0}{[P]_0}} \right] \quad [III]$$

Dans un régime d'échange rapide, on peut considérer que  $[PL]/[P]_0$  est directement proportionnel aux différences des PDCs ( $\Delta\delta$  (ppm)). L'équation devient donc;

$$\Delta\delta = \frac{\Delta\delta_{max}}{2} \left[ 1 + \frac{[L]_0}{[P]_0} + \frac{K_D}{[P]_0} - \sqrt{\left( 1 + \frac{[L]_0}{[P]_0} + \frac{K_D}{[P]_0} \right)^2 - 4 \frac{[L]_0}{[P]_0}} \right] \quad [IV]$$



Sachant que les variables  $[L]_0$ ,  $[P]_0$ ,  $\Delta\delta_{\max}$  et  $\Delta\delta$  peuvent toutes être déterminées lors des expériences de titrage RMN, cette équation permet de déterminer la valeur du  $K_d$ . Par contre, il est important de noter que selon cette équation, plus la valeur de  $K_D$  est élevée, plus il est difficile d'atteindre la saturation ( $\Delta\delta_{\max}$ ) (**Figure 1.17**).



**Figure 1.17. Simulation de courbes pour quantifier le  $K_D$  à partir des PDC.**

Graphique des courbes théoriques montrant la dépendance de la forme liée d'une protéine en fonction du ratio ligand sur protéine ( $L_0/P_0$ ) pour différentes constantes de dissociation  $K_D$ . Figure tiré de [125].

#### **1.5.4 Déterminer la structure d'une interaction entre protéines à l'aide de la RMN**

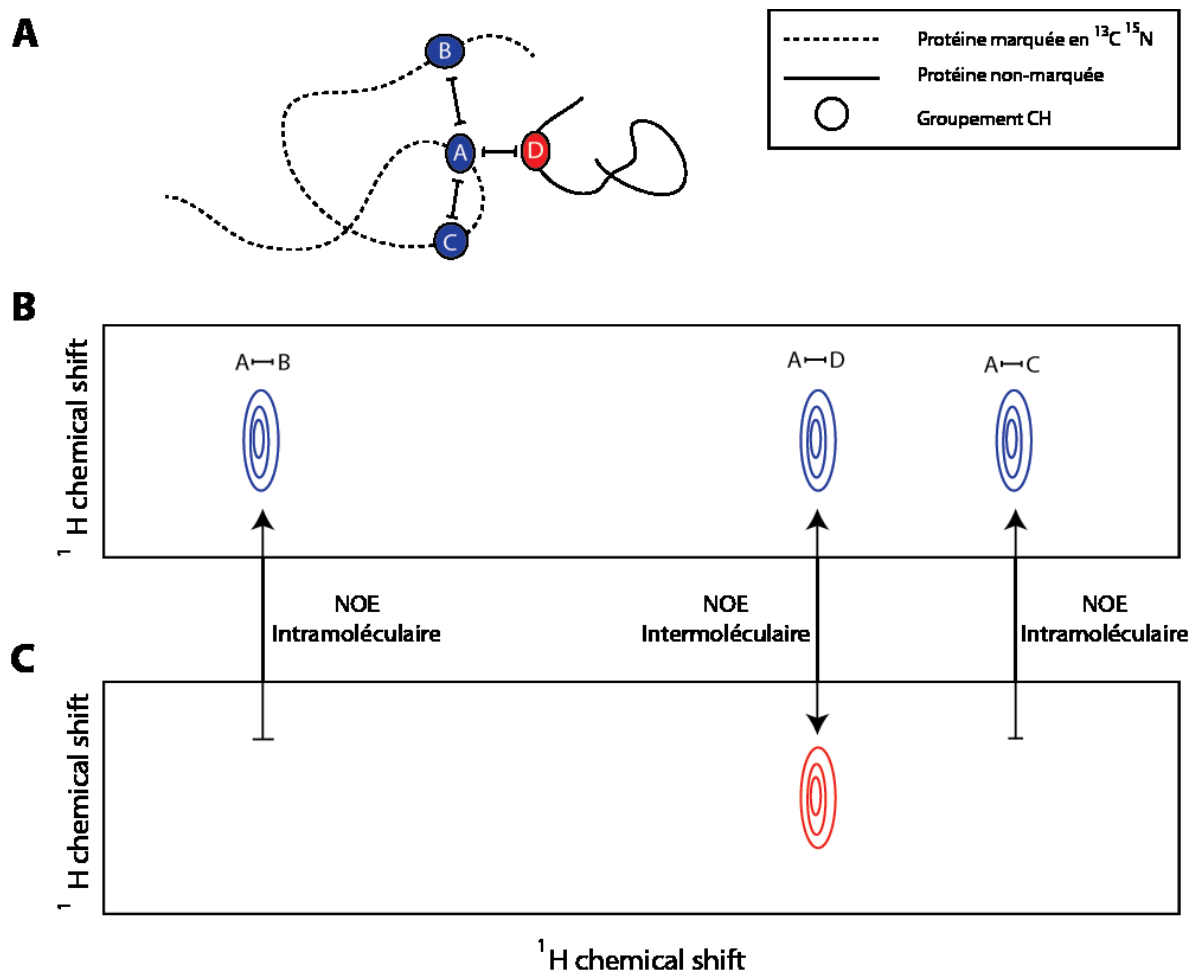
Bien que le titrage RMN soit une manière simple et efficace pour analyser les liaisons entre protéines, une compréhension détaillée de l'interaction ne peut être obtenue qu'avec la structure 3D à haute résolution du complexe. Celle-ci aide à la compréhension des fonctions biologiques du complexe et par exemple, d'obtenir les informations nécessaires pour être capable de développer des inhibiteurs spécifiques à l'interface de l'interaction. De plus, la structure 3D d'un complexe est requise pour la classification et la compréhension mécanistique des domaines fonctionnels. La

détermination de structures par RMN peut se décrire en deux étapes principales : la collection des contraintes structurales et le calcul de structure.

#### 1.5.4.1 *Collection de contraintes structurales*

Obtenir une structure par RMN est un long procédé qui implique l'acquisition de différentes expériences permettant d'obtenir les contraintes structurales d'un complexe. Deux expériences essentielles pour obtenir la structure d'une interaction entre protéines sont  $^{15}\text{N}$  et  $^{13}\text{C}$  NOESY (Nuclear overhauser effect spectroscopy). Basé sur l'effet nucléaire overhauser qui se définit comme étant une interaction entre deux spins de noyaux atomiques à travers l'espace, ces deux expériences permettent de corréler l'environnement spatial de chaque lien C-H et N-H avec tous leurs protons voisins situés à moins de 5 Å; ce qui représente la majorité des atomes d'une protéine. L'intensité du signal observé est inversement proportionnelle à la distance qui sépare les voisins et ainsi, permet d'extraire des milliers de contraintes de distances qui définissent l'ensemble de la structure.

Lors de la détermination de la structure d'un complexe entre deux protéines, les spectres NOESY sont enregistrés avec une seule des deux protéines marquée afin d'éviter le chevauchement et l'empilement des signaux. Dans ces spectres, on peut observer les NOE entre les liens C-H ou N-H de la molécule marquée avec les hydrogènes  $^1\text{H}$  de cette même molécule (distance intramoléculaire), mais, aussi avec ceux de la molécule non-marquée (distance intermoléculaire). Ainsi, il est difficile de déterminer quel signal correspond à une distance entre protons intramoléculaires ou intermoléculaires. Un spectre NOESY intermoléculaire peut être enregistré pour discriminer les signaux NOE provenant de la molécule marquée de ceux de la molécule non-marquée. Ce filtre de NOE facilite l'analyse des spectres NOESY et révèle les atomes impliqués dans la formation du complexe sous forme de distances intermoléculaires (**Figure 1.18**). Ensemble, les distances tirées des spectres de  $^{15}\text{N}$  NOESY,  $^{13}\text{C}$  NOESY et NOESY intermoléculaire seront utilisées pour déterminer la structure à haute résolution du complexe.



**Figure 1.18. Schéma du principe des spectres NOESY.**

(A) Représentation de différents signaux NOE observables à partir du groupement CH de A d'une protéine marquée. (B) Portion 2D du spectre 3D provenant de l'expérience  $^{13}\text{C}$ -edited HMQC-NOESY théorique de la région du groupement CH de A de la protéine marquée où les signaux intra et intermoléculaire sont visibles. (C) Portion 2D du spectre NOESY intermoléculaire théorique de la région du groupement CH de A de la protéine marquée où seuls les signaux intermoléculaires sont visibles.

#### 1.5.4.2 Calcul et validation de la structure

Après avoir obtenu les contraintes structurales vient l'étape du calcul de structure. Le calcul de structure en RMN consiste à trouver les structures de plus basse énergie respectant à la fois les contraintes expérimentales et les contraintes physico-chimiques des acides aminés. Chaque violation de contrainte engendre une accumulation d'énergie. Pour minimiser celle-ci, la plupart des logiciels utilisent le recuit simulé (mieux connu sous le nom de « simulated annealing »), qui est une méthode d'exploration de l'espace conformationnelle. Ces calculs proposent des centaines de structures répondant au

maximum de contraintes qui sont ensuite triées en fonction de leur énergie. Celles possédant la plus basse énergie sont choisies et constituent l'ensemble de structures répondant aux contraintes expérimentales fournies au logiciel.

L'étape finale pour obtenir la structure d'un complexe à partir de la RMN consiste à valider la qualité du modèle calculé par le logiciel sans tenir compte des données enregistrées. Dans ce processus les angles de torsion de la chaîne principale sont d'abord vérifiés par le diagramme de Ramachandran pour valider la structure secondaire du modèle calculé. Ensuite, les longueurs des liaisons et les angles entre les atomes sont également vérifiés par rapport à leur valeur idéale. Finalement, la précision des structures est mesurée en calculant l'écart quadratique moyen (Root Mean Square Deviation : RMSD) entre les coordonnées atomiques des différentes structures finales. Ceci nous permet de savoir à quel point les structures calculées sont semblables entre elles. Des structures désordonnées ou divergentes auront un RMSD plus élevé que les structures ordonnées ou convergentes.

#### **1.5.5 Analyse des structures à l'aide d'études fonctionnelles.**

Suite à l'obtention de la structure 3D du complexe d'intérêt, les protéines impliquées sont sujettes à des analyses fonctionnelles. Il n'y a pas de méthode standard pour analyser une protéine de manière fonctionnelle puisque le type d'expériences dépend entièrement des protéines étudiées et des conclusions biologiques visées. Dans notre cas, pour vérifier nos hypothèses portant sur le rôle que peut avoir un UBD sur un TAD, nous avons élaboré des essais de luciférase pour déterminer l'effet de mutations spécifiques sur la stabilité des AT.

Cette expérience consiste à déterminer l'effet d'une mutation sur la stabilité d'une protéine par rapport à sa forme sauvage. Dans cet essai, la stabilité de la protéine est déterminée à partir de sa quantité présente dans des cellules humaines 24h après leur transformation par un plasmide codant pour cette protéine fusionnée à la Renilla luciférase (Rluc) qui sert de rapporteur. Cette dernière émet, de manière proportionnelle

à sa concentration, une longueur d'onde détectable suite au contact avec son substrat. L'expérience est répétée avec des quantités croissantes de plasmides protéine-RLuc où le signal devrait augmenter linéairement avec la quantité de plasmides insérés. La pente de cette régression linéaire permet de déterminer la dépendance de l'expression selon la stabilité de la protéine insérée. Théoriquement, une protéine moins stable devrait avoir une valeur expression/stabilité plus élevée qu'une protéine stable. Ainsi, la comparaison entre les résultats obtenus avec un plasmide protéine-RLuc sauvage et une forme mutée de cette protéine permet de déterminer l'effet de la mutation sur sa stabilité.

Cependant, l'efficacité de transformation de cellules humaines peut être variable. Ainsi, le vecteur contrôle exprimant le rapporteur Firefly luciférase (FLuc), qui absorbe et émet à des longueurs d'ondes différentes de la RLuc, est inséré en même temps à concentration constante. Le ratio RLuc/FLuc sert à normaliser les erreurs qui peuvent survenir à cause de l'efficacité imparfaite de la transfection et de la viabilité des cellules. Aussi, afin de garder la même quantité de plasmides insérés, un plasmide vide est ajouté pour garder la quantité d'ADN insérée constante pour chaque expérience. En somme, cette expérience permet de déterminer si les mutations qui empêchent la formation du complexe entre les DAD d'AT et l'UBI ont un effet sur la stabilité de ces activateurs.



## **2.Caractérisation d'une interaction entre le TAD d'EKLF et l'ubiquitine**





## 2.1. Mise en contexte de l'article.

Notre laboratoire possède une très grande expertise sur les TAD acides et leur effet sur la transcription, notamment avec la caractérisation structurale et fonctionnelle des complexes entre la sous-unité Tfb1/p62 de FTIIH et les TAD acides de VP16 et de p53 [8, 120]. La prémisses du laboratoire était fait à partir d'un alignement de séquence qui démontrait la forte similarité entre plusieurs TAD acides humains comme EKLF, p53 et viraux comme VP16, EBNA2, HPV-1 et NSs. Durant le début de ma maîtrise, j'ai participé à mettre en place des protocoles de production et de purification pour des nouveaux TAD acides tels qu'EKLF, SREBP1a, Gal4, EBNA2 et HPV-1 afin de pouvoir déterminer leurs similitudes avec les TAD acides déjà testés. Ces efforts ont mené à la publication de deux articles qui caractérisaient le complexe entre la sous-unité de Tfb1/p62 de FTIIH et les TAD acides d'EKLF [16] et de la protéine virale EBNA2 [127]. Le protocole de production et de purification mis en place fonctionne pour la majorité des TAD acides de l'alignement de séquences. Ainsi, ceci nous a permis de fabriquer une banque de TAD acides pour le laboratoire afin de pouvoir tester plusieurs hypothèses à leur sujet. Pour le projet de ma maîtrise, j'ai enquêté sur la forte ressemblance entre les domaines DAD des TAD acides et les domaines UBD.

Dans l'article qui suit, nous voulions répondre principalement à la question : **est-ce que les domaines DAD des TAD acides sont capables d'interagir avec l'UBI de façon non covalente pour réguler leur fonction paradoxale d'activation dépendante de leur dégradation?** Dans ce manuscrit, nous avons décrit le tout premier cas d'un DAD capable de lier l'UBI et montré que cette interaction était importante pour sa fonction. Malgré leur ressemblance avec des UBD, ce ne sont pas tous les DAD testés qui sont capables de lier l'UBI. Parmi ceux qui étaient capables d'interagir, seul EKLF avait été étudié au préalable dans notre laboratoire. Ainsi, à des fins de comparaison, l'article qui suit présente une étude détaillée sur la structure du complexe entre le DAD d'EKLF (référé comme EKLFTAD1 dans l'article) et l'UBI.

## 2.2. Contribution des auteurs

**Dr Luca Raiola** a réalisé et analysé la majeure partie des expériences présentes et a participé à la rédaction de l'article.

**Mathieu Lussier-Price** a participé à la production des protéines et de leurs mutants utilisés pour toutes les expériences *in vitro*. Il a aussi aidé à l'analyse des spectres RMN et a participé à la rédaction de l'article.

**Dr David Gagnon** a participé aux expériences de dégradation *in vivo*.

**Dr Julien Lafrance-Vanasse** a réalisé une partie de la collecte des spectres RMN et a participé à l'analyse des spectres.

**Dr Xavier Mascle** a participé aux expériences de dégradation *in vivo*.

**Dr Pascale Legault** a conseillé et vérifié le déroulement du projet ainsi que révisé l'article.

**Dr Jacques Archambault** a conseillé et vérifié le déroulement du projet ainsi que révisé l'article.

**Dr James G. Omichinski** a dirigé le projet, participé à la résolution de structure et rédigé l'article avec le Dr Luca Raiola et Mathieu Lussier-Price.

### 2.3. The acidic TAD of EKLF directly binds ubiquitin.

#### **Structural Characterization of a Non-covalent Complex Between Ubiquitin and the Acidic Transactivation Domain of the Erythroid Specific Factor EKLF**

*Luca Raiola<sup>1</sup>, Mathieu Lussier-Price<sup>1</sup>, David Gagnon<sup>2</sup>, Julien Lafrance-Vanasse<sup>1</sup>, Xavier Mascle<sup>1</sup>, Genevieve Arseneault<sup>1</sup>, Pascale Legault<sup>1</sup>, Jacques Archambault<sup>1,2</sup> and James G. Omichinski<sup>1\*</sup>*

[REDACTED]

[REDACTED]

[REDACTED]

Running Title: *The acidic TAD of EKLF directly binds ubiquitin.*

## ***SUMMARY***

Like several other acidic transactivation domains (TAD), the minimal TAD from the erythroid-specific transcription factor EKLF (EKLFTAD) has been shown to contribute both to its transcriptional activity as well as to its ubiquitin(UBI)-mediated degradation. In this manuscript, we examine the activation-degradation role of the acidic TAD of EKLF and demonstrate that the first forty residues (EKLFTAD1) within this region form a direct non-covalent interaction with UBI. NMR structural studies of an EKLFTAD1-UBI complex show that EKLFTAD1 adopts a 14-residue  $\alpha$ -helix that forms the recognition interface with UBI in a similar manner as the UBI-interacting helix of Rabex5. We also identify a similar interaction between UBI and the activation-degradation region of SREBP1a, but not with the activation-degradation regions of p53, GAL4 and VP16. These results suggest that select activation-degradation regions like the ones found in EKLF and SREBP1a function in part through their ability to form non-covalent interactions with UBI.

## **INTRODUCTION**

Ubiquitin (UBI) is a highly conserved protein that is found in all eukaryotic organisms (Hershko et al., 2000). As a post-translational modification, UBI is covalently attached to the  $\epsilon$ NH<sub>2</sub> group of lysine residues in target proteins via an enzymatic pathway that includes an E1-activating enzyme, an E2-conjugating enzyme and an E3-ligase (Hochstrasser, 2000; Pickart, 2001; Weissman, 2001). Historically, post-translational modifications with UBI were considered to regulate protein levels by targeting them for degradation. In protein degradation, target proteins are modified with UBI or with poly-UBI chains containing four or more UBIs branching from lysine 48 (K48) and this leads to their recognition and proteolysis by the 26S proteasome (Elsasser and Finley, 2005; Miller and Gordon, 2005). Over the last several years, it has been shown that post-translational modification with either UBI or various forms of poly-UBI at one of several different lysine residues on UBI can play important roles in regulating a number of other cellular functions in addition to degradation (Geng et al., 2012; Haglund and Dikic, 2005; Ulrich, 2012).

One important feature that enables UBI to play important roles in cellular regulation is its ability to function in non-covalent protein-protein interaction through interactions with proteins containing UBI-binding domains (UBD) (Dikic et al., 2009; Hurley et al., 2006; Winget and Mayor, 2010). Given the ubiquitous occurrence of UBI throughout cells, a number of different UBD have been identified (for a review see (Dikic et al., 2009)). Several of these UBDs contain helices that form the interface through recognition of a hydrophobic patch formed by Leu8, Ile44 and Val70 of UBI (Dikic et al., 2009; Hurley et al., 2006; Winget and Mayor, 2010). As a protein interaction module, UBI can interact with UBDs either as free UBI or as a fusion protein. Since UBI can itself be poly-ubiquitinated at numerous different lysine residues (6, 11, 23, 27, 29, 33, 48 and 63), it has been proposed that by using different combinations of poly-UBI branching patterns the potential number of functions regulated through interactions with UBI is virtually unlimited (Ikeda and Dikic, 2008) and UBI has been shown to regulate cell signaling pathways, DNA repair pathways, and transcription (Geng et al., 2012; Haglund and Dikic, 2005; Ulrich, 2012).

The role that UBI plays in transcriptional regulation has been extensively investigated over the last several years (Geng et al., 2012). Based on these studies, it is now evident that UBI functions in a number of different manners to regulate transcription. Initial studies demonstrated that poly-ubiquitination and degradation through the ubiquitin-proteasome system (UPS) functioned as a straightforward process for tightly regulating levels of transcription factors (Kubbutat et al., 1997). However, the crucial role of UBI in transcription regulation became more evident when a direct correlation was established between the ability of many proteins to activate transcription and their post-translational modification with UBI (Molinari et al., 1999). These studies established that several transcriptional activators contained a “degron region” that targets them for destruction via the UPS, and that this “degron region” overlapped with their transcriptional activation domain (TAD) (Salghetti et al., 2001; Salghetti et al., 2000). Additional studies indicated that the function of the overlapping activation-degradation region in regulating transcription is extremely complex involving a number of different mechanisms depending on the activator involved (Geng et al., 2012). For example, acidic TADs from p53, VP16 (Herpes Simplex Viral protein 16) and GAL4 appear to directly associate with a specific UBI E3-ligase, and this leads directly to their ubiquitination, activation and degradation (Archer et al., 2008; Haupt et al., 1997; Kubbutat et al., 1997; Le Cam et al., 2006; Muratani et al., 2005; Salghetti et al., 2001). In the case of SREBP1a (Sterol Regulatory Enhancer Binding Protein 1a), both the acidic TAD and the DNA-binding domain function as “degron regions”, and binding of the E3-ligase Fbw7 (F-Box and W-repeat domain-containing 7) occurs subsequent to a DNA-binding dependent phosphorylation reaction (Bengoechea-Alonso and Ericsson, 2009; Sundqvist et al., 2005; Sundqvist and Ericsson, 2003). The results from these studies have led to a number of different hypotheses for explaining the overlapping activation/degradation function of TADs, and it is now clear that different mechanisms are most likely involved depending on the pathway (Geng et al., 2012).

Like a number of other factors that regulate hematopoietic development, it has been previously shown that the Erythroid-Krüppel like factor (EKLF) is post-translationally

modified by UBI on several lysine residues, and UBI-mediated degradation plays an important role in regulating its turnover (Quadrini and Bieker, 2006). EKLF (also referred to as KLF1) is a member of the Krüppel-like protein family that is characterized by the presence of three Cys<sup>2</sup>His<sup>2</sup> zinc-finger domains at the extreme carboxy-terminus of the protein (Miller and Bieker, 1993). Functionally, EKLF is an erythroid-specific transcription factor that plays a key role in red blood cell development, in part by enhancing  $\beta$ -globin gene expression through binding to CACCC-consensus sequences located in the promoter region of the gene (Geng et al., 2012). Mutations in either the zinc-finger domains of EKLF or the CACCC sequences in the  $\beta$ -globin promoter regions result in a decrease in  $\beta$ -globin expression underlying a number of blood cell related diseases (Feng et al., 1994; Perkins et al., 1995). In addition to the three zinc-finger domains, EKLF contains a unique bipartite activation that includes a minimal TAD located in the amino-terminal 90 amino acids (Chen and Bieker, 1996). This minimal TAD of EKLF can be further subdivided into two independent domains EKLFTAD1 (residues 1-40) and EKLFTAD2 (residues 51-90) that are both capable of independently activating transcription (Chen and Bieker, 1996). EKLFTAD1 and EKLFTAD2 are both highly acidic and contain a PEST (Sequence rich in proline, glutamic acid, serine and threonine) region (Quadrini and Bieker, 2006). Recent studies have demonstrated that deletion of either PEST regions, especially the PEST1 region within EKLFTAD1, significantly retards ubiquitin-mediated degradation of EKLF without lowering the overall ubiquitination of EKLF (Quadrini and Bieker, 2006). These results suggest that the minimal TAD of EKLF has an overlapping activation/degradation function that is critical to the role of EKLF in red blood cell development.

Although it has been clearly demonstrated that EKLF is degraded in a ubiquitin-dependent manner that requires the PEST1 region within the minimal TAD (Quadrini and Bieker, 2006), there is very little known regarding the function the TAD of EKLF plays in its UBI-mediated degradation. In this manuscript, we establish that UBI forms a direct non-covalent interaction with the minimal TAD of EKLF that requires residues within the PEST1 region of EKLFTAD1. NMR structural characterization of an EKLFTAD1-UBI complex demonstrates that a 14-residue amphipathic helix in EKLFTAD1 forms the recognition

interface with UBI through interactions with a hydrophobic patch formed by Leu8, Ile44 and Val70. The structure of EKLFTAD1 in complex with UBI indicates that the helix is strikingly similar to the helix of the Motif-Interacting with Ubiquitin (MIU) domain of Rabex-5 bound to UBI (Lee et al., 2006; Penengo et al., 2006). NMR titration studies with other overlapping activation-degradation regions from p53, GAL4, VP16 and SREBP1a indicates that binding of activation-degradation regions to UBI is not a universal attribute, since only the region from SREBP1a appears to form a non-covalent interaction with UBI. Using a cell-based assay, we also establish that, like for SREBP1a, the DNA-binding domain of EKLF also plays a significant role in regulating its stability. Together, these results suggest that the activation-degradation regions of EKLF and SREBP1a function in similar manners through non-covalent interactions with UBI and that their role in degradation is distinct from the role of similar activation-degradation regions of p53, VP16 and GAL4, which function by directly binding to UBI E3-ligases.

## **RESULTS**

### *EKLFTAD1 Binds Non-covalently to UBI*

Although earlier studies indicated that deletion of either of the two PEST regions significantly retards EKLF degradation, it appeared that the PEST1 region (residues 32-52 in human EKLF) plays a more significant role than the PEST2 region (residues 54-67) (Quadrini and Bieker, 2006). Interestingly, both EKLFTAD1 (residues 1-40) and EKLFTAD2 (residues 50-90) are composed predominantly of acidic and hydrophobic amino acid residues and the spacing pattern of the hydrophobic residues is surprisingly similar to what is observed in the sequences of known Ubiquitin Interacting Motifs (UIMs) and Motifs Interacting with Ubiquitin (MIUs) (Dikic et al., 2009). Given this sequence similarity between EKLFTAD1/EKLFTAD2 and UIMs/MIUs and the fact that the minimal EKLFTAD functions as an activation-degradation region, we were interested to determine if the minimal EKLFTAD is capable of forming a direct non-covalent protein-protein interaction with UBI. To test the direct interaction of the EKLFTAD with UBI, NMR chemical shift perturbation studies were performed (**Figure 2. 1** and **Supplementary Figure 2.S 1**). In



these experiments, the addition of unlabeled UBI to either  $^{15}\text{N}$ -labeled EKLFTAD1 (**Figure 1A and 1C**),  $^{15}\text{N}$ -labeled EKLFTAD2 (**Supplementary Figure 2.S 1**) or  $^{15}\text{N}$ -labeled EKLFTAD (**Figure 2.1B and 1D**) causes significant changes in both the  $^1\text{H}$  and  $^{15}\text{N}$  chemical shifts for several signals in the  $^1\text{H}$ - $^{15}\text{N}$ -HSQC spectrums. In the titration with the full minimal EKLFTAD, the signals that undergo the most significant chemical shift changes correspond to residues in the EKLFTAD1 subdomain (**Figure 2.1**). In contrast, the signals that correspond to residues in the EKLFTAD2 subdomain undergo only very minor chemical shift changes in the titration (**Supplementary Figure 2.S1**). Quantitative analysis of the chemical shift changes occurring during the titrations indicates that EKLFTAD1 and the EKLFTAD2 subdomain bind to UBI with apparent dissociation constants ( $K_D$ ) of 300  $\mu\text{M}$  and 1.5 mM, respectively. Taken together, the results indicate that the minimum EKLFTAD binds directly to UBI and that residues in the EKLFTAD1 subdomain form the interface with UBI with similar affinity as measured for other UIMs/MIUs (Lee et al., 2006; Penengo et al., 2006; Swanson et al., 2003).

#### *UBI Interacts Non-covalently with EKLFTAD1*

Several structures of UBI in non-covalent complexes with different binding partners demonstrate that the most common interaction interface of UBI is a hydrophobic patch formed by Leu8-Ile44-Val70 (Dikic et al., 2009; Hurley et al., 2006; Winget and Mayor, 2010). To determine if EKLFTAD1 also interacts with this hydrophobic patch, NMR chemical shift perturbation studies were performed with unlabeled EKLFTAD1 and either  $^{15}\text{N}$ -labeled UBI or  $^{15}\text{N}/^{13}\text{C}$ -labeled UBI. Addition of unlabeled EKLFTAD1 to  $^{15}\text{N}$ -labeled UBI causes significant changes in the  $^1\text{H}$  and  $^{15}\text{N}$  chemical shifts for several signals in the  $^1\text{H}$ - $^{15}\text{N}$ -HSQC spectrum of UBI (**Figure 2.2A**). Likewise, addition of unlabeled EKLFTAD1 to  $^{15}\text{N}/^{13}\text{C}$ -labeled UBI causes significant changes in the  $^1\text{H}$  and  $^{13}\text{C}$  chemical shifts for several signals in the  $^1\text{H}$ - $^{13}\text{C}$ -HSQC spectrum of UBI (**Figure 2.2B**). Consistent with what is observed with UIMs/MIUs binding to UBI (Dikic et al., 2009; Hurley et al., 2006; Winget and Mayor, 2010), the most significant chemical shift changes in the  $^1\text{H}$ - $^{15}\text{N}$ -HSQC spectrum of  $^{15}\text{N}$ -labeled UBI are centered around three distinct regions between Thr7-Thr14 (L1 and  $\beta$ 2), Leu43-Glu51 ( $\beta$ 3, L3 and  $\beta$ 4) and Leu67-Leu73 ( $\beta$ 5 and C-terminus) (**Figure 2.2C-D**). The

role of the Leu8-Ile44-Val70 patch is further supported by the fact that significant  $^1\text{H}$  and  $^{13}\text{C}$  chemical shift changes in the  $^1\text{H}$ - $^{13}\text{C}$  HSQC spectrum are observed for the signals of the Hy1 and Hy2 of Ile44, the Hy1 and Hy2 of Val70 and the H $\delta$ 1 and H $\delta$ 2 of Leu8 upon addition of EKLFTAD1 (**Figure 2.2B**).

#### *NMR Structure Determination of the UBI-EKLFTAD1 Complex*

Given that EKLFTAD1 binds to the Leu8-Ile44-Val70 patch of UBI, the NMR solution structure of the EKLFTAD1-UBI complex was determined in order to compare it with structures of other UIMs/MIUs in complex with UBI. The NMR structure of the EKLFTAD1-UBI complex is well defined by the NMR data (**Table II**). The 20 lowest-energy structures are characterized by good backbone geometry, no significant restraint violation, and low pairwise r.m.s.d. values (**Figure 2.3A**). The structures were calculated using a total of 1792 NOE-derived distance restraints including 15 intermolecular NOEs, 39 hydrogen-bond restraints and 100 dihedral angle restraints. A total of 100 structures were calculated and all of them satisfy the experimental constraints with no NOE violation greater than 0.2 Å and no backbone angle violation greater than 2° (**Table II**).

#### *Overall Structure of the EKLFTAD1-UBI Complex*

The structure of UBI in complex with EKLFTAD1 is a typical UBI fold that consists of a 12-residue alpha helix (H1), a short 3(10)-helix (H2) and a five-stranded mixed  $\beta$ -sheet ( $\beta$ 1- $\beta$ 5) (Vijaykumar et al., 1985) (**Figure 2.3A-B**). EKLFTAD1 is disordered in the complex with the exception of a 14-residue amphipathic  $\alpha$ -helix that forms between residues Thr23 and Ala36 upon binding UBI (**Figure 2.3B**). Consistent with the chemical shift mapping results, the EKLFTAD1 helix forms an interface with UBI on a surface that includes L1,  $\beta$ 3, L3,  $\beta$ 5 and the C-terminus that are all within the Leu8-Ile44-Val70 patch. Likewise, the signals for residues that form the alpha helix in EKLFTAD1 undergo the most significant chemical shift changes when UBI is added. With the exception of the 14-residue alpha helix, EKLFTAD1 is disordered and this transition of a short region from a disordered to an ordered conformation is consistent with what has been observed for several other acidic

TADs upon binding to other target proteins (Di Lello et al., 2006; Ferreon et al., 2009; Langlois et al., 2008; Wojciak et al., 2009).

As described above, the interface of the EKLFTAD1-UBI complex is predominantly formed through interactions involving the amphipathic  $\alpha$  helix of EKLFTAD1 and the Leu8-Ile44-Val70 hydrophobic patch of UBI. Like the interaction of the MIU of Rabex5 with UBI (Lee et al., 2006; Penengo et al., 2006), the helix of EKLFTAD1 binds in the reverse orientation as to what is observed with UIMs binding to UBI (Swanson et al., 2003) (**Figure 4**). Three hydrophobic residues from the hydrophobic face of the helix of EKLFTAD1 (Phe27, Leu28 and Trp31) appear along the recognition interface, and Leu28 and Trp31 appear to contribute significantly to the interaction with UBI (**Figure 2.4A**). In the complex, the side chain of Leu28 is inserted into a pocket formed by Lys6, Leu8, Ile44, His68 and Val70 of UBI where it is in position to form methyl-methyl contacts with the side chains of Ile44 and Val70 of UBI (**Figure 2.4B**). Trp31 is the second key residue of EKLFTAD1 at the recognition interface with UBI. The aromatic side chain of Trp31 inserts into a shallow pocket that is formed by Leu8, Arg42, Ile44, Gln49, Val70 and Arg72 of UBI (**Figure 2.4C**), where it forms several van der Waals contacts with methyl groups of Leu8, Ile44 and Val70 from UBI. In this orientation, the indole ring of Trp31 is also in position to form a cationic- $\pi$  interaction with the side chain of Arg42 from UBI. In the case of Phe28 of UBI, it appears to play a minor role in the binding to UBI through an interaction between its aromatic side chain and the H $\gamma$  methyl group of Ile44.

In addition to hydrophobic interactions, the negatively charged residues from EKLFTAD1 are in position to form several potential ionic interactions with positive charges on UBI in the NMR structures. In particular, Asp25 at the N-terminal end of the  $\alpha$  helix of EKLFTAD1 is in position to form a salt bridge with Lys6 of UBI (**Figure 2.4B and Supplementary Figure 2.S2**). Likewise, Glu34 and Glu35 at the C-terminal end of the  $\alpha$  helix of EKLFTAD1 are in position to form ionic interactions with a positively charged surface formed by Arg42, Arg72 and Arg74 of UBI (**Figure 2.4C and Supplementary Figure 2.S2**). These potential

electrostatic interactions at both ends of the EKLFTAD1 helix function to establish the orientation of the  $\alpha$  helix on the UBI surface.

#### *Mutational Studies of the EKLFTAD1 Interaction with UBI*

The NMR solution structure of the EKLFTAD1-UBI complex demonstrates that the recognition interface is formed through interactions between the hydrophobic face of the EKLFTAD1 amphipathic  $\alpha$ -helix and UBI (**Figures 2.3 and 2.4**). To assess the relative contributions of key residues of the EKLFTAD1  $\alpha$ -helix, we performed mutagenesis at several positions. Three EKLFTAD1 point mutants (F27S, L28S, and W31S) were generated, and the binding of the variant proteins to UBI was tested by NMR chemical shift perturbation experiments. In each experiment, unlabeled UBI was incrementally added to the  $^{15}\text{N}$ -labeled EKLFTAD1 variants as described in the previous section. During these three titrations, we were not able to reach saturation or determine values for the dissociation constants. Furthermore, after addition of UBI to variant (F27S, L28S and Trp31S) EKLFTAD1 proteins, the resulting  $^1\text{H}$ - $^{15}\text{N}$  HSQC spectrums were very similar to that of their free form (**Supplementary Figure 2.S3**). These results confirm that Phe27, Leu28 and Trp31 on the hydrophobic face of the EKLFTAD1  $\alpha$ -helix are crucial for binding to UBI. The structure of the EKLFTAD1-UBI complex also indicates that the Leu8-Ile44-Val70 hydrophobic patch of UBI establishes the recognition interface with EKLFTAD1. To examine the contribution of this hydrophobic patch, we used NMR chemical shift perturbations experiments to test the binding of an I44A variant of UBI (UBI44A), since this variation has been shown to disrupt the binding of UIMs/MIUs to UBI (Shih et al., 2002). In this experiment, unlabeled UBI44A was incrementally added to the  $^{15}\text{N}$ -labeled EKLFTAD1. As observed with UIMs binding to UBI, addition of the unlabeled I44A variant to  $^{15}\text{N}$ -labeled EKLFTAD1 fails to induce significant changes in either the  $^1\text{H}$  or  $^{15}\text{N}$  chemical shifts of the 2D  $^1\text{H}$ - $^{15}\text{N}$ -HSQC spectrum of EKLFTAD1 (**Supplementary Figure 2.S4**). Taken together, these experiments with variant proteins clearly demonstrate that the non-covalent interaction between EKLFTAD1 and UBI is very specific.

*The Role of the UBI-interacting region of EKLFTAD1 in EKLF Stability.*

Previous studies have shown that EKLF is degraded in a ubiquitin-mediated manner and that the two PEST domains within the minimal TAD play an important function in this pathway. In an attempt to more quantitatively assess the role of the UBI-interacting domain from EKLF in regulating its stability, we employed a luciferase-based stability assay where variant EKLF proteins are expressed in fusion with Renilla Luciferase (RLuc) and compared the steady-state level of wild-type EKLF (Mesplede et al., 2012). Three EKLF constructs were tested with RLuc fused to the C-terminus; the first construct encoded for full-length human EKLF (EKLF-RLuc), the second encoded for EKLF with a deletion of the full minimal TAD ( $\Delta$ TAD-EKLF-RLuc) and the third encoded for EKLF with serine substituted for F27, L28 and W31 (F27S/L28S/W31S-EKLF). Increasing concentrations of the EKLF-RLuc,  $\Delta$ TAD-EKLF-RLuc or F27S/L28S/W31S-EKLF vectors were transfected into human C33a cells and luciferase levels were measured 24 h after transfection. As an internal control, a constant amount of a firefly luciferase (FLuc) expression vector was used to normalize for transfection efficiency and cell viability. As expected, the RLuc/FLuc ratio of each construct was found to increase linearly as a function of the 5-different concentrations of EKLF expression vector tested (**Figure 2.5A**). Linear regression analysis of this data was used to calculate a slope indicative of the stability of each EKLF constructs. Under these assay conditions, the levels of the EKLF-RLuc protein was slightly more than 2-fold lower than both the  $\Delta$ TAD-EKLF-RLuc and the F27S/L28S/W31S-EKLF proteins (**Figure 2.5B**). In addition, the levels of the TAD-EKLF-RLuc and the F27S/L28S/W31S-EKLF proteins were very similar. These results are in agreement with earlier studies showing that the PEST1 region within the minimal TAD of EKLF plays a significant role in ubiquitin-mediated degradation of EKLF (Quadrini and Bieker, 2006) and further establishes that the UBI-interacting domain in EKLF plays an important role in regulating its steady-state levels.

*The Activation-Degradation Region of SREBP1a Binds UBI*

Given that EKLFTAD1 can bind directly to UBI in a manner similar to the MIU of Rabex-5 (Lee et al., 2006; Penengo et al., 2006), we were interested to determine if other

well-characterized activation-degradation regions can also form a direct non-covalent interaction with UBI (Salghetti et al., 2000). To test this, NMR chemical shift perturbation studies were performed with activation-degradation regions from the TAD of p53 (p53TAD1; residues 1-40), SREBP-1a (SREBP-1aTAD; residues 1-50), VP16 (VP16N; residues 412-456) and GAL4 (GAL4TAD; residues 857-881) (Salghetti et al., 2000; Sundqvist and Ericsson, 2003) (**Figure 2.6 and Supplementary Figure 2.S5**). These four TADs were chosen because their sequences all contain a region with significant homology to the region of EKLFTAD1 that forms the recognition interface with UBI (**Supplementary Figure 2.S6**). The four TADs were  $^{15}\text{N}$ -labeled and their binding to unlabeled UBI was monitored by NMR chemical shift perturbation experiments. As observed with EKLFTAD1 (**Figure 2.1A**), the sequential addition of unlabeled UBI to  $^{15}\text{N}$ -labeled SREBP-1aTAD causes significant changes in both the  $^1\text{H}$  and  $^{15}\text{N}$  chemical shifts for several signals of SREBP-1aTAD in the 2D  $^1\text{H}$ - $^{15}\text{N}$ -HSQC spectra (**Figure 2.6**). In contrast, addition of unlabeled UBI to either  $^{15}\text{N}$ -labeled p53TAD1, VP16N or GAL4TAD did not induce any significant changes in their signals in the 2D  $^1\text{H}$ - $^{15}\text{N}$ -HSQC spectra (**Supplementary Figure 2.S5**). Based on these results, it appears that only select acidic TADs that function as activation-degradation regions have the capacity to bind non-covalently to UBI as part of their function and this includes SREBP1a and EKLFTAD1.

## **DISCUSSION**

Hematopoietic development requires that cell-specific factors participate in a complex network of macromolecular interactions in order to activate lineage-specific expression programs while simultaneously repressing expression of alternative blood lineages. Given the complexity of the process, precise functioning levels of hematopoietic specific transcription factors are required and this can be achieved through different mechanisms including controlling rates of gene expression and/or rates of proteolytic degradation. In this manuscript, we examine the mechanism of ubiquitin-mediated degradation of the erythroid specific transcription factor EKLF. Previous studies have shown that EKLF levels are predominantly regulated by UBI-mediated proteolysis (Quadrini and Bieker, 2006). In addition, a PEST region located within the minimal TAD of

EKLF plays a significant role in its stability suggesting that this domain has both activation and degradation functions (Quadrini and Bieker, 2006). We verify that the minimal TAD of EKLF is essential for its degradation and demonstrate that EKLFTAD1 within this region forms a direct non-covalent protein-protein interaction with UBI. NMR structural studies of the EKLFTAD1-UBI complex show that EKLFTAD1 adopts a 14-residue amphipathic  $\alpha$ -helix when bound to UBI, and this helix interacts with UBI in a manner similar to the MUI domain of Rabex5 (Lee et al., 2006; Penengo et al., 2006). Studies with TADs from several other proteins demonstrate that binding to UBI is not a universal attribute for activation-degradation regions, since only the activation-degradation regions of EKLF SREBP1a appear to form a direct interaction with UBI.

The NMR structure of the EKLFTAD1-UBI complex demonstrates that EKLFTAD1 binds along the Leu8-Ile44-Val70 hydrophobic patch of UBI in a similar manner as the MUI of Rabex5 (Lee et al., 2006; Penengo et al., 2006). The structures of EKLFTAD1-UBI and Rabex5 MUI-UBI are very similar to the NMR structure of the UIM of Vsp27 bound to UBI with the notable exception that the  $\alpha$  helix of Vsp27-UIM binds UBI in a reverse orientations with respect to its N- and C-termini (Lee et al., 2006; Penengo et al., 2006; Swanson et al., 2003). Therefore, it is interesting to compare the EKLF-UBI structure with both the Rabex5-UBI structure and the Vsp27-UBI structure. Aligning the sequences from the UBI-interacting helices of EKLF, Rabex5 and Vsp27 (**Figure 2.7A**), it appears there is a consensus motif of X-X-a- $\phi$ - $\phi$ -X-X- $\phi$ -X-X-a-a where “a” is an acidic residue (Asp or Glu) and “ $\phi$ ” is a hydrophobic residue. In addition, the superposition of the three interacting helices in complex with UBI indicates that they are very similar to each other (**Figure 2.7B**). For all three UBI-interacting helices, the hydrophobic residue in position five of the helix makes direct contacts with the Leu8-Ile44-Val70 hydrophobic patch of UBI. In the case of Rabex5 and Vsp27, this residue is an Ala (Ala58 and Ala266 respectively), whereas in EKLF this corresponds to Leu28. The superposition of the EKLF-UBI and Rabex5-UBI structures demonstrates that one of the methyl groups of Leu28 from EKLF is in a similar orientation along the surface of UBI as the methyl group of Ala58 from Rabex5 (**Figure 2.7C**). In both cases, the methyl groups are in position to form methyl-methyl contacts with the side

chains of Ile44 and Val70 of UBI. The other key residue for the interaction of EKLF, Rabex5 and Vps27 helices with UBI is the hydrophobic residue in position eight of the helix. This residue corresponds to Leu61 in Rabex5, Ile263 in Vps27 and Trp31 in EKLF and in all cases it is inserted into a large pocket formed by Leu8, Arg42, Ile44, Gln49, Val70 and Arg72. Superimposing the EKLF-UBI and Rabex5-UBI structures shows that the side chains of Leu61 of Rabex5 and Trp31 of EKLF are positioned in almost the same location on the surface of UBI (**Figure 2.7D**). In comparison with the UBI-interacting helices of Rabex5 and Vps27, the  $\alpha$ -helix of EKLF is slightly shifted relative to UBI due to the presence of the bulkier Trp31 side chain (**Figure 2.7B**). This small shift allows for the correct positioning of the aromatic ring of Trp31 as well as the other methyl group of Leu28 in the Leu8-Ile44-Val70 hydrophobic patch of UBI.

A number of mechanisms have been proposed to account for the overlapping activation/degradation functions associated with the acidic TAD of several transcriptional regulatory proteins (Bhat and Greer, 2011; Geng et al., 2012; Kodadek, 2010; Kodadek et al., 2006). Our results here suggest that the activation-degradation regions of EKLF and SREBP1a may function in similar manners in part through their ability to form a direct non-covalent protein-protein interaction with UBI. This mechanism would be distinct from what is observed for the activation-degradation regions of p53, VP16 and Gal4, which do not appear to form a non-covalent interaction with UBI. The activation-degradation regions of p53, VP16 and Gal4 have all been shown to form direct interactions with their respective UBI E3-ligases. In all three cases, the binding to the E3 ligase enhances ubiquitination, activation and degradation via the UPS (Archer et al., 2008; Haupt et al., 1997; Kubbutat et al., 1997; Le Cam et al., 2006; Muratani et al., 2005; Salghetti et al., 2001). For both SREBP1a and EKLF, the role of their acidic N-terminal activation-degradation regions appears to be more complex. In the case of SREBP1a, it must first bind to DNA to initiate the activation/degradation process (Punga et al., 2006). Once bound to DNA, SREBP1a is phosphorylated on a number of serine residues near its C terminus, and this leads to the recruitment of its E3 ligase Fbw7 (Bengoechea-Alonso and Ericsson, 2009; Punga et al., 2006; Sundqvist et al., 2005). EKLF does not appear to contain



a sequence similar to the phosphorylation region found in SREBP1a, but a recent study demonstrated that the PEST1 region in EKLAD forms a direct interaction with the Ppm1b phosphatase (Yien and Bieker, 2012). Currently, the E3 ligase or ligases responsible for EKLF ubiquitination have not yet been identified, although it is known that EKLF is ubiquitinated at several lysine residues and deletion of the PEST1 region appears to result in a slight hyper-ubiquitination (Quadrini and Bieker, 2006).

In conclusion, the precise role of the activation-degradation regions of EKLF in regulating its activation and degradation functions is clearly complex, but it appears that its ability to form non-covalent interactions with UBI may be essential to these roles. For both EKLF and SREBP1a, deletion studies demonstrate that these activation-degradation regions are not essential for their ubiquitination (Quadrini and Bieker, 2006; Sundqvist and Ericsson, 2003). This would indicate that these regions do not function by stabilizing an interaction with an UBI-charged E2-conjugating enzyme that leads to mono-ubiquitination independent of an E3 ligase as has been demonstrated for other proteins containing UBDs (Hoeller et al., 2007). One possible explanation is that the activation-degradation regions of EKLF and SREBP1a function by forming either an interaction with ubiquitinated transcriptional regulatory proteins (Geng et al., 2012) or an intermolecular interaction with one of their own ubiquitinated lysine residues (Hoeller et al., 2006) that is important both for their activation and degradation functions. Such non-covalent binding to ubiquitinated proteins has been suggested to be the functional role of the MUI domain of Rabex5 (Mattera and Bonifacino, 2008). However, additional studies are needed to more clearly define the role of the non-covalent interaction between UBI and the activation-degradation regions of EKLF and SREBP1a in regulating their transcriptional activity and degradation.

## **EXPERIMENTAL PROCEDURES**

### *Cloning of EKLF-RLuc Constructs.*

The plasmid for expressing EKLF-RLuc was prepared starting with the cDNA of human EKLF provided by Dr. James Bieker (Mount Sinai Hospital, New York, NY). This was cloned into pRLuc in order to express full-length EKLF (residues 1-363) as a fusion protein

with Renilla luciferase at the C terminus of the protein. The  $\Delta$ TAD-EKLF-RLuc (EKLF residues 91-363 fused to RLuc) and F27S/L28S/W31S-EKLF-RLuc (EKLF with serine substitutions at Phe27, Leu28 and Trp31 fused to RLuc) expression plasmids were prepared by deletion or mutation of the EKLF-RLuc plasmid. All plasmids were sequenced to confirm their identity.

#### *EKLF Stability Assay.*

Increasing amounts of plasmids encoding EKLF-RLuc,  $\Delta$ TAD-EKLF-RLuc or F27S/L28S/W31S-EKLF-RLuc were co-transfected with a constant amount of FLuc normalization plasmid (2.5 ng) into  $2.5 \times 10^4$  C33a cells in white 96-well plates. For each transfection, the total amount of transfected DNA was kept constant by the addition of the appropriate empty vector. Twenty-four hours after transfection, FLuc and RLuc activities were measured using the Dual-Glo luciferase system (Promega) and a GloMax 96-well luminometer (Promega) according to the instructions from the manufacturer. For each construct, 5 different amounts of expression vector (6.25, 12.5, 25, 50, 100 ng) were tested in triplicates and in two separate experiments. The error bars represent standard deviation.

#### *Statistical Analysis.*

For the EKLF stability assay, linear regression analysis of all 30 values was used to calculate a slope diagnostic of the expression/stability of each EKLF construct. Standard deviations from quadruplicate values are indicated as error bars.

#### *Cloning, Expression and Purification of Recombinant Proteins.*

The EKLFTAD (residues 1-90 of hEKLF), EKLFTAD1 (residues 1-40 of hEKLF), EKLFTAD (residues 50-90 of hEKLF), p53TAD1 (residues 1-40 of human p53), VP16N (residues 412-456 of Herpes Simplex Virus protein 16), SREBP1aTAD (residues 1-50 of human SREBP1a) and Gal4TAD (residues 851-881 of GAL4) proteins were all cloned into the pGEX-2T vector (GE Healthcare). Point mutants of EKLFTAD1 were generated using the QuikChange Mutagenesis Kit (Stratagene). All proteins were expressed as GST-fusion

proteins in E.coli host strain TOPP2, bound to GSH resin (GE Healthcare), cleaved with thrombin and purified over a Q-sepharose column as previously described (Di Lello et al., 2005). UBI was purified as described (Sass et al., 1999). The UBI44A protein (UBI with I44A substitution) was generated using the QuikChange Mutagenesis Kit and purified as described for UBI.  $^{15}\text{N}$ -labeled and  $^{15}\text{N}/^{13}\text{C}$ -labeled proteins were prepared in media containing  $^{15}\text{NH}_4\text{Cl}$  and/or  $^{13}\text{C}_6$ -glucose as the sole nitrogen and carbon sources.

#### *NMR Samples.*

NMR samples for titrations consisted in 0.4 mM of  $^{15}\text{N}$ -UBI in 20 mM sodium phosphate buffer (pH 6.5) [90 %  $\text{H}_2\text{O}/10\%$   $\text{D}_2\text{O}$ ] and unlabeled proteins (EKLFTAD, EKLFTAD1, EKLFTAD2) were added in stepwise increments. The same experimental conditions were used for titration of the  $^{15}\text{N}$ -labeled TAD proteins with either unlabeled UBI or UBI44A. For NMR structural studies of the EKLFTAD1-UBI complex, four samples were prepared in 20 mM sodium phosphate buffer (pH 6.5) [90% $\text{H}_2\text{O}/10\%$   $\text{D}_2\text{O}$  or 100 %  $\text{D}_2\text{O}$ ] : 0.8 mM of either  $^{15}\text{N}$ -labeled UBI and 4.0 mM unlabeled EKLFTAD1, 0.8 mM  $^{15}\text{N}/^{13}\text{C}$ -labeled UBI and 4.0 mM unlabeled EKLFTAD1, 0.8 mM  $^{15}\text{N}$ -labeled EKLFTAD1 and 4.0 mM of unlabeled UBI, 0.8 mM  $^{15}\text{N}/^{13}\text{C}$ -labeled EKLFTAD1 and 4.0 mM of unlabeled UBI.

#### *NMR Spectroscopy.*

NMR spectra were collected at 300 K on Varian Unity Inova 500, 600 and 800 MHz spectrometers. The backbone and aliphatic side chain resonances ( $^1\text{H}$ ,  $^{15}\text{N}$  and  $^{13}\text{C}$ ) were assigned by a series of triple resonance experiments [HNCO (Kay et al., 1994), HNCACB (Wittekind and Mueller, 1993), (HB)CBCA(CO)NNH (Grzesiek and Bax, 1992), C(CO)NNH (Logan et al., 1992), H(CCO)NNH (Logan et al., 1992), and HCCH-COSY (Ikura et al., 1991). Distance restraints were obtained from  $^{15}\text{N}$ -edited NOESY-HSQC and  $^{13}\text{C}$ -edited NOESY-HSQC experiments (90 ms mixing time). Intermolecular NOEs were obtained from a 3D  $^{15}\text{N}$ ,  $^{13}\text{C}$  (F1)-filtered, (F3)-NOESY-HSQC (Zwahlen et al., 1997). NMR data were processed by NMRPipe/NMRDraw (Delaglio et al., 1995) and analyzed with ccpnmr Analysis (Vranken et al., 2005).

*Structure Calculations.*

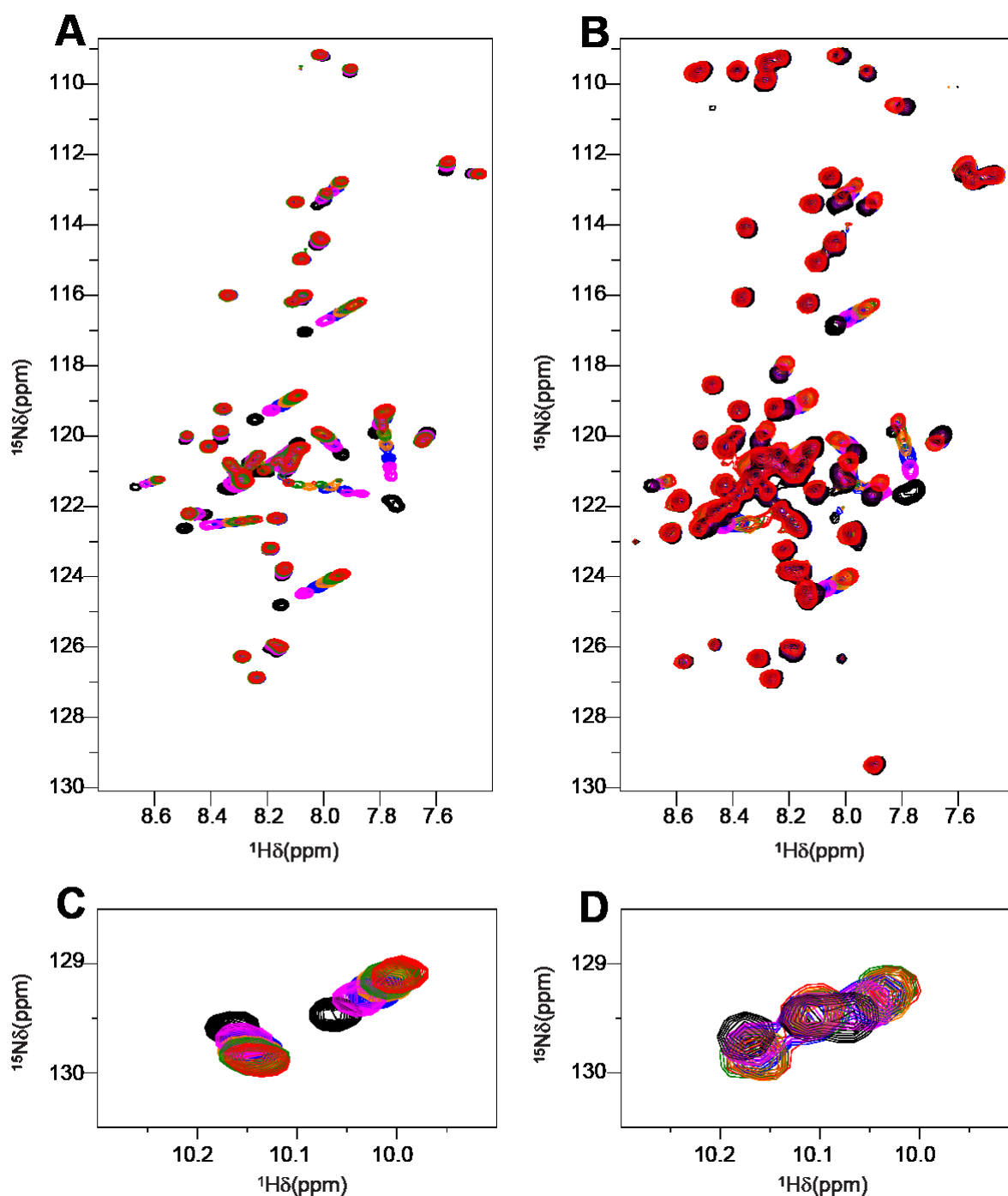
Structure calculations were performed with the program CNS using a combination of torsion angle dynamics (TAD) and Cartesian dynamics (Brunger et al., 1998). The backbone dihedral angles were obtained using the program TALOS (Cornilescu et al., 1999). Starting from an extended structure with standard geometry, 50 conformers were calculated satisfying all the experimental constraints with no NOE violation greater than 0.2 Å and no dihedral angle violations greater than 2°. The quality of the structures was analyzed using PROCHECK-NMR (Laskowski et al., 1996) and MOLMOL (Koradi et al., 1996). All of the figures representing the structures were generated with the program PyMol.

Accession Codes. Protein Data Bank : 2MBH. BMRB : 19399

**ACKNOWLEDGEMENTS**

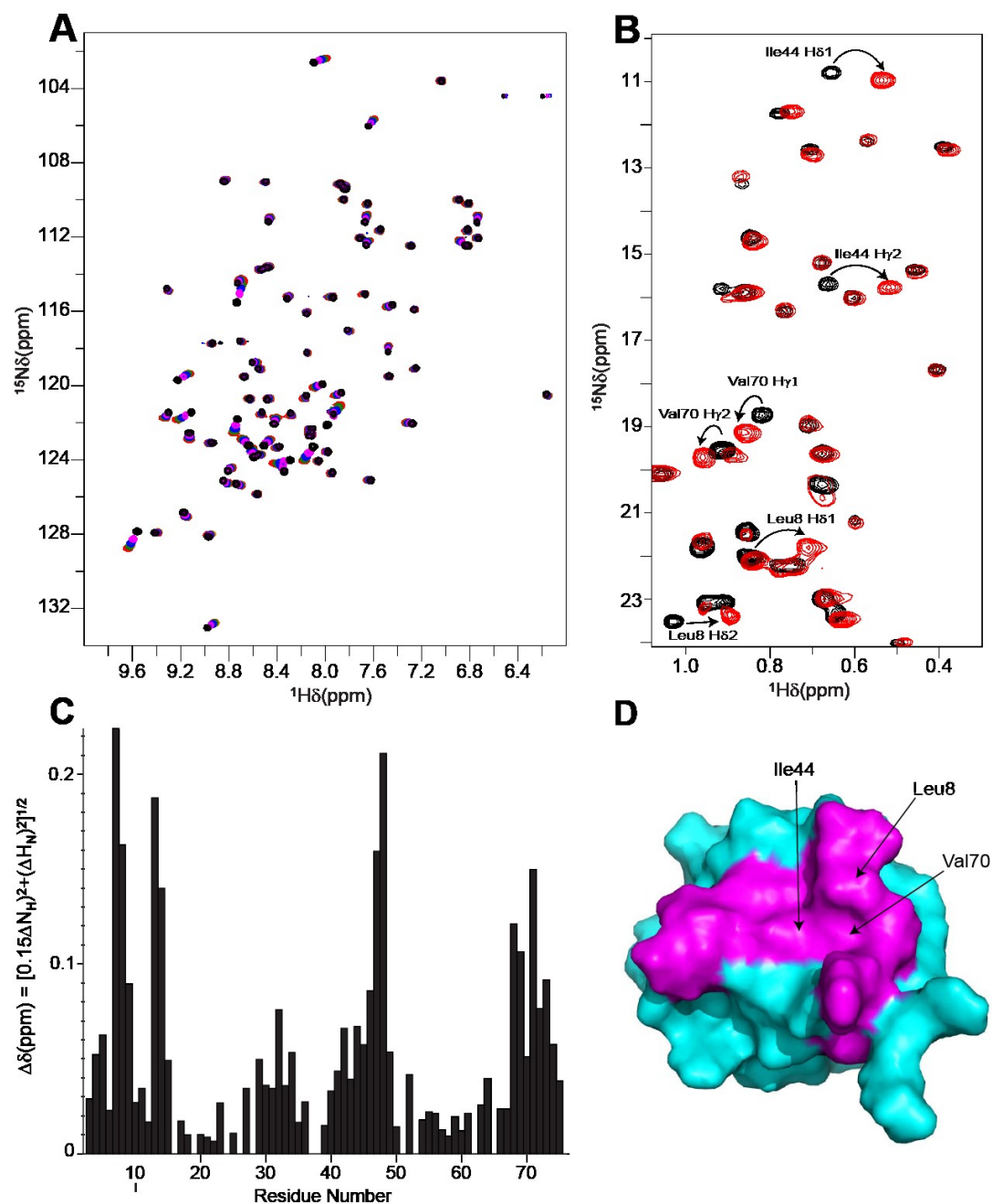
This work was supported by grants from the Canadian Institutes for Health Research (CIHR) to JGO (MOP-74739), the Canadian Cancer Society Research Institute (CCSRI) to JA and The Cancer Research Society to JA. P.L. is a Canadian Research Chair in Structural Biology and Engineering of RNA. 800 MHz NMR experiments were recorded at the Québec/Eastern Canada High Field NMR Facility, supported by the Natural Sciences and Engineering Research Council of Canada. The microcalorimeter was purchased with funds from the NSERC.

## FIGURES



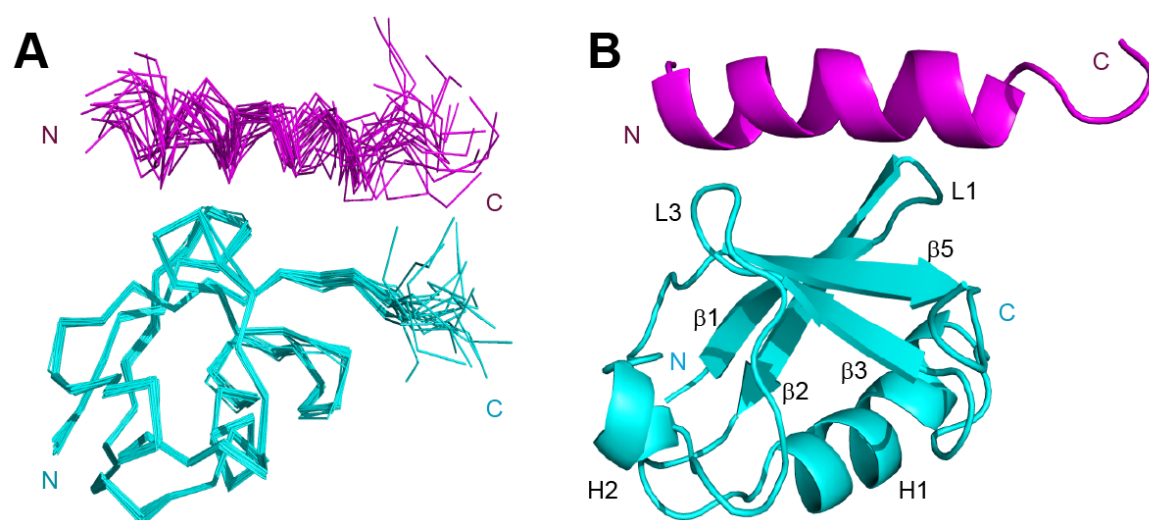
**Figure 2. 1. The TAD of EKL binds to UBI.**

Overlay of the 2D  $^1\text{H}$ - $^{15}\text{N}$  HSQC spectra of  $^{15}\text{N}$ -labeled EKLFTAD1 (A) and  $^{15}\text{N}$ -labeled EKLFTAD (B) in the free form (black) and in the presence of 1 (pink), 2 (blue), 3 (orange), 4 (green) and 5 (red) molar equivalents of UBI. The side chain of Trp (H $\epsilon$ 1) from the titrations with  $^{15}\text{N}$ -labeled EKLFTAD1 and  $^{15}\text{N}$ -labeled EKLFTAD are shown in (C) and (D) respectively. See also **Supplementary Figure 2.S 1**.



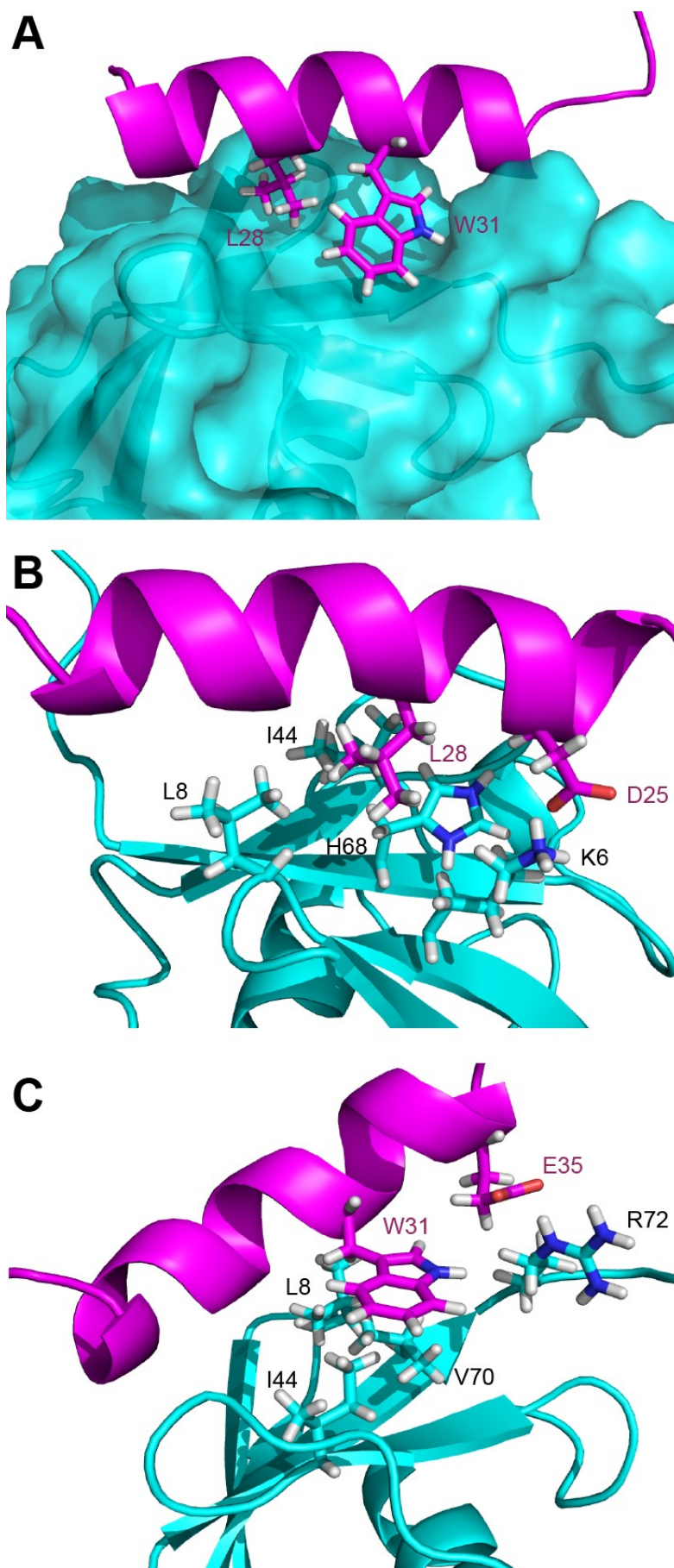
**Figure 2. EKLFTAD1 binds the Leu8-Ile44-Val70 hydrophobic patch of UBI.**

(A) Overlay of the 2D  $^1\text{H}$ - $^{15}\text{N}$  HSQC spectra of  $^{15}\text{N}$ -labeled UBI in the free form (black) and in the presence of 1 (pink), 2 (blue), 3 (orange), 4 (green) and 5 (red) molar equivalents of EKLFTAD1. (B) Overlay of the methyl region from 2D  $^1\text{H}$ - $^{13}\text{C}$  HSQC spectra of  $^{15}\text{N}/^{13}\text{C}$ -labeled UBI in the free form (black) and in the presence of 5 (red) molar equivalents of EKLFTAD1. Black arrows indicate signals from methyl groups located in the Leu8-Ile44-Val70 patch of UBI. (C) Histogram of the variation in chemical shift  $[\Delta\delta(\text{ppm})] = [0.15\Delta N_H]^2 + (\Delta N_H)^2]^{1/2}$  observed in the 2D  $^1\text{H}$ - $^{15}\text{N}$  HSQC spectrum of UBI upon addition of EKLFTAD1. (D) Ribbon model (cyan) of free UBI (PDB Accession Code : 1UBQ). The residues showing significant changes in chemical shifts in the 2D  $^1\text{H}$ - $^{15}\text{N}$  HSQC and 2D  $^1\text{H}$ - $^{13}\text{C}$  HSQC of  $^{15}\text{N}$ -labeled and  $^{15}\text{N}/^{13}\text{C}$ -labeled UBI upon formation of the EKLFTAD1-UBI complex are highlighted (magenta).



**Figure 2. 3. Structure of the EKLFTAD1-UBI complex.**

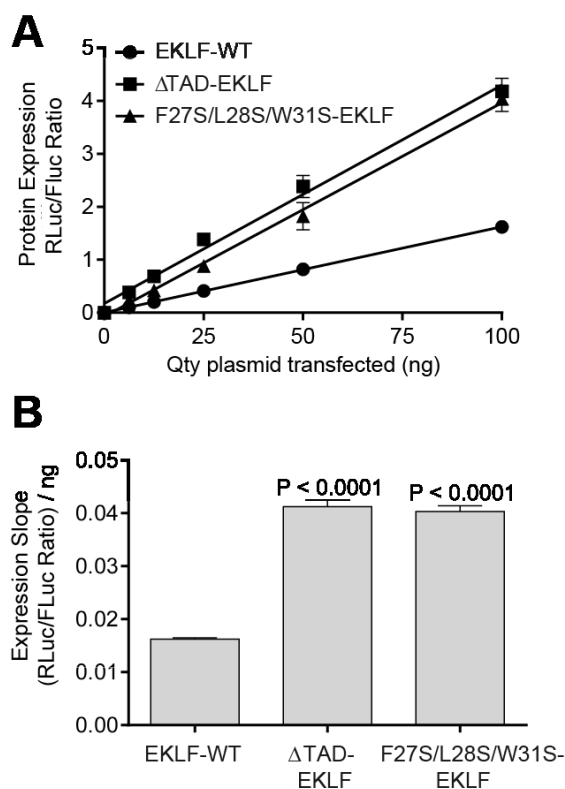
(A) Overlay of the twenty lowest-energy structures of the complex between UBI (cyan) and the UBI-interacting helix of EKLFTAD1 (magenta). The structures were superimposed using backbone atoms C', C $\alpha$  and N of residues 2-74 of UBI and residues 23-36 of EKLFTAD1. (B) Ribbon model of the lowest-energy conformer of the complex between UBI (cyan) and EKLFTAD1 (magenta).





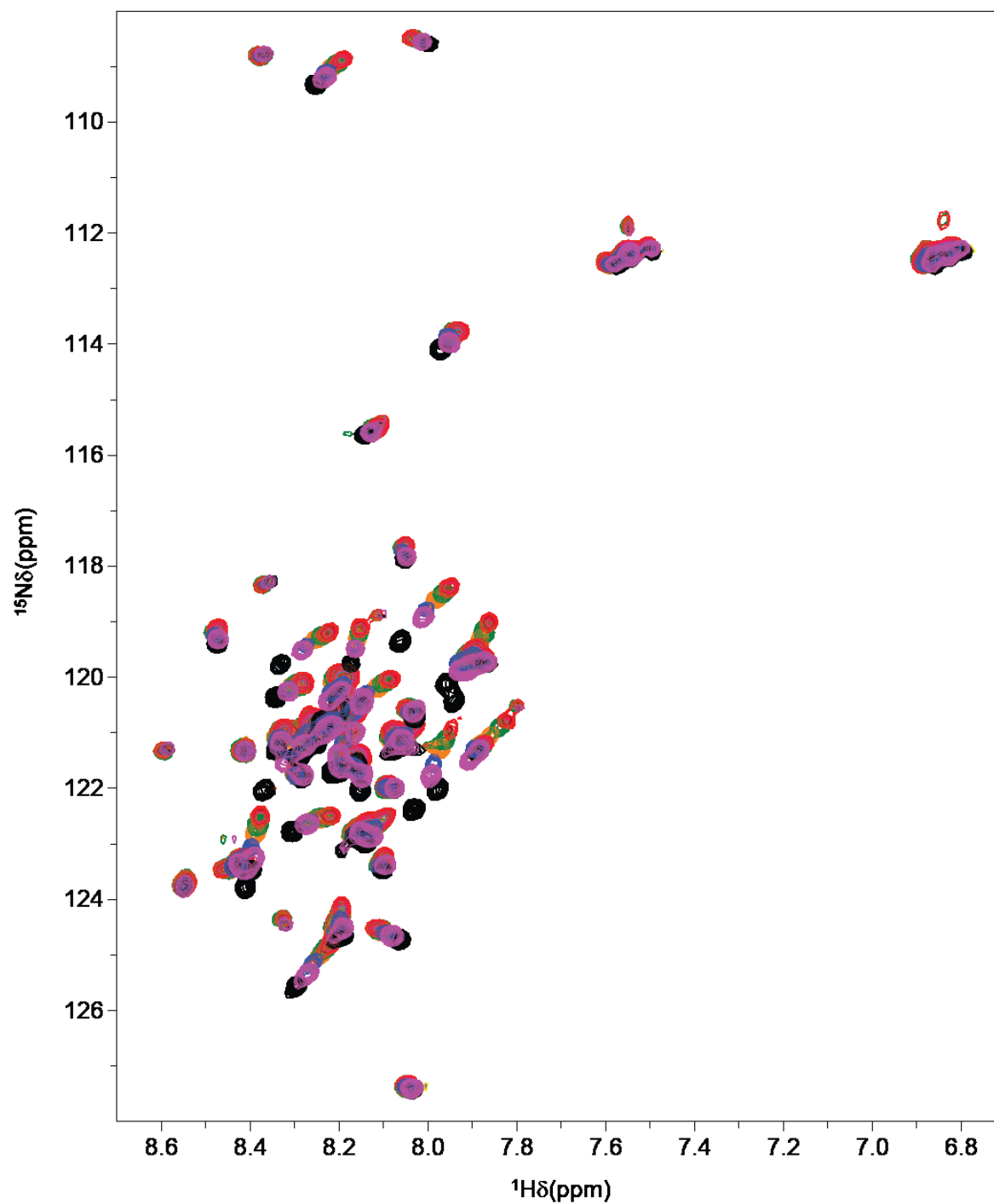
#### Figure 2. 4. Structural details of the EKLFTAD1-UBI interface.

In (A), the backbone of EKLFTAD1 (magenta) is represented as a ribbon whereas UBI is shown as a molecular surface (cyan). The side chains of Phe27, Leu28 and Trp31 from the hydrophobic face of EKLFTAD1  $\alpha$ -helix are shown as sticks. Trp31 of EKLFTAD1 is shown inserted into a shallow pocket formed by Leu8, Arg42, Ile44, Gln49, Val70 and Arg72 of UBI and Leu28 of EKLFTAD1 is shown inserted into a pocket formed by Lys6, Leu8, Ile44, His68 and Val70 of UBI. In (B-C), the backbone of EKLFTAD1 and UBI are represented as ribbons. The side chains of Lys6, Leu8, Ile44 and Val70 of UBI as well as Asp25 and Leu28 of EKLFTAD1 are shown as sticks in (B). The side chain of Leu28 of EKLFTAD1 is in position to form methyl-methyl contacts with the side chains of Leu8 and Ile44 of UBI. In addition, the carboxyl group of Asp25 from EKLFTAD1 is in position to form a salt bridge with the amino group from the side chain of Lys6 of UBI. In (C), the side chains of residues Leu8, Ile44, Val70 and Arg72 of UBI and Phe27, Trp31 and Glu35 of EKLFTAD1 are shown as sticks. The indole ring of Trp31 of EKLFTAD1 interacts with methyl groups of Leu8, Ile44 and Val70 of UBI and the carboxyl group of Glu35 of EKLFTAD1 is in position to form an ionic interaction with the guanidino group of Arg72 of UBI. See also **Supplementary Figure 2.S 2** and **Supplementary Figure 2.S 4**.



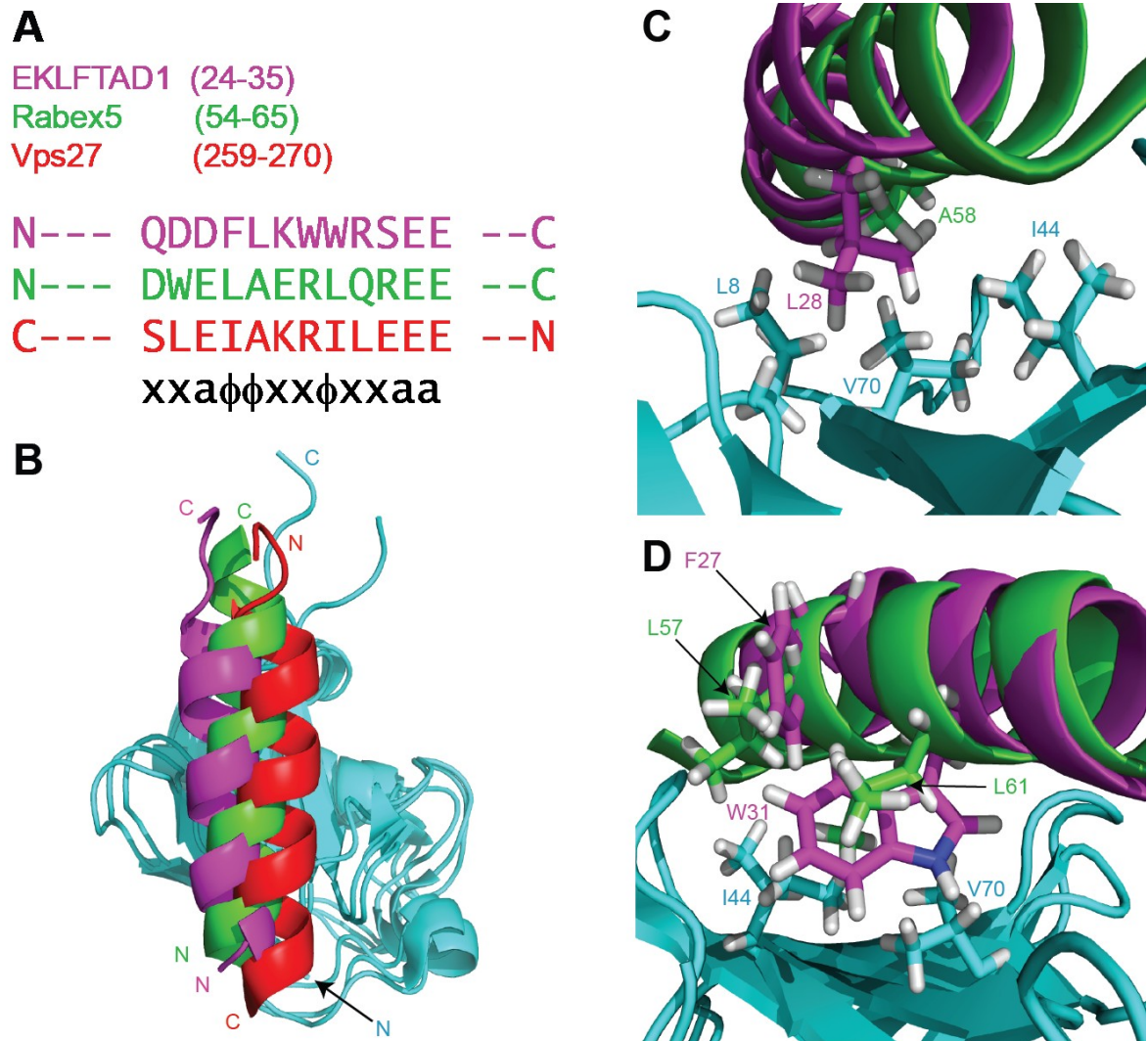
#### Figure 2. 5. The minimal TAD of EKLF functions as a “degron” in cells.

(A) Luciferase activity measured in C33A cells transfected with increasing amounts of either the EKLF-RLuc,  $\Delta$ TAD-EKLF-RLuc or F27S/L28S/W31S-EKLF-RLuc expression vector. For each measurement, the levels of RLuc activity were normalized to those of FLuc expressed from a co-transfected plasmid. (B) Bar graph showing the expression slopes of EKLF,  $\Delta$ TAD-EKLF and F27S/L28S/W31S-EKLF calculated from the data in (A). Standard deviations from quadruplicate values are indicated as error bars, but in some cases are too small to be visible.



**Figure 2. 6. The acidic ADR of SREBP-1a binds to UBI.**

Overlay of the 2D  $^1\text{H}$ - $^{15}\text{N}$  HSQC spectra of  $^{15}\text{N}$ -labeled SREBP-1aTAD in the free form (black) or in the presence of 1 (pink), 2 (blue), 3 (orange), 4 (green) and 5 (red) molar equivalents of UBI. See also **Supplementary Figure 2.S 5** and **Supplementary Figure 2.S 6**.



**Figure 2. 7. Sequence and structure comparison of the UBI-interacting helices from EKLFTAD1, Rabex5 and Vps27.**

(A) Sequence alignment of the UBI-interacting helix of EKLFTAD1 with that of the MUI of Rabex5 and UIM of Vps27. Note that the Vps27 sequence is aligned in the inverse orientation relative to EKLFTAD1 and Rabex5 in accordance with its inverse orientation in complex with UBI. In the consensus sequence, “ $\phi$ ” represents a hydrophobic amino acid residue, “a” an acidic amino acid residue and “x” any amino acid residue. (B) Overlay of the ribbon structures of EKLFTAD1 (magenta), Rabex5 (green; PDB Accession Code : 2C7N) and Vps27 (red; PDB Accession Code : 1Q0W) in complex with UBI (cyan). The  $\alpha$ -helices of EKLFTAD1 and Rabex5 are in the same orientation (N-terminus to C-terminus), whereas the  $\alpha$ -helix of Vps27 is in the reverse orientation. (C-D) Overlay of the ribbon structures of UBI (cyan) and the UBI-interacting helices of EKLFTAD1 (magenta) and Rabex5 (green). The side chains of Leu28 from EKLFTAD1 and Ala58 from Rabex5 are shown as sticks inserting in a similar manner into a pocket formed by Ile44 and Val70 of UBI in (c). In (d), the side chains of Phe27 and Trp31 from EKLFTAD1 and Leu57 and Leu61 from Rabex5 are shown as sticks inserting into a pocket formed by Leu8, Ile44 and Val70 of UBI in a similar manner.

## REFERENCES

- Archer, C.T., Delahodde, A., Gonzalez, F., Johnston, S.A., and Kodadek, T. (2008). Activation domain-dependent monoubiquitylation of Gal4 protein is essential for promoter binding in vivo. *J. Biol. Chem.* 283, 12614-12623.
- Bengoechea-Alonso, M.T., and Ericsson, J. (2009). A phosphorylation cascade controls the degradation of active SREBP1. *J. Biol. Chem.* 284, 5885-5895.
- Bhat, K.P., and Greer, S.F. (2011). Proteolytic and non-proteolytic roles of ubiquitin and the ubiquitin proteasome system in transcriptional regulation. *Biochim. Biophys. Acta-Gene Reg. Mech.* 1809, 150-155.
- Brunger, A.T., Adams, P.D., Clore, G.M., Gros, P., Grosse-Kunstleve, R.W., Jiang, J.-S., Kuszewski, J., Nilges, M., Pannu, N.S., Read, R.J., et al. (1998). Crystallography and NMR system (CNS): A new software system for macromolecular structure determination. *Acta Crystallogr. D* 54, 905-921.
- Chen, X.Y., and Bieker, J.J. (1996). Erythroid Kruppel-like factor (EKLF) contains a multifunctional transcriptional activation domain important for inter- and intramolecular interactions. *EMBO J.* 15, 5888-5896.
- Cornilescu, G., Delaglio, F., and Bax, A. (1999). Protein backbone angle restraints from searching a database for chemical shift and sequence homology. *J. Biomol. NMR* 13, 289-302.
- Delaglio, F., Grzesiek, S., Vuister, G.W., Zhu, G., Pfeifer, J., and Bax, A. (1995). NMRPipe: A multidimensional spectral processing system based on UNIX pipes. *J. Biomol. NMR* 6, 277-293.
- Di Lello, P., Jenkins, L.M.M., Jones, T.N., Nguyen, B.D., Hara, T., Yamaguchi, H., Dikeakos, J.D., Appella, E., Legault, P., and Omichinski, J.G. (2006). Structure of the Tfb1/p53 complex: insights into the interaction between the p62/Tfb1 subunit of TFIIH and the activation domain of p53. *Mol. Cell* 22, 731-740.
- Di Lello, P., Nguyen, B.D., Jones, T.N., Potempa, K., Kobor, M.S., Legault, P., and Omichinski, J.G. (2005). NMR structure of the amino-terminal domain from the Tfb1 subunit of TFIIH and characterization of its phosphoinositide and VP16 binding sites. *Biochemistry* 44, 7678-7686.
- Dikic, I., Wakatsuki, S., and Walters, K.J. (2009). Ubiquitin-binding domains - from structures to functions. *Nature reviews Mol. Cell. Biol.* 10, 659-671.
- Elsasser, S., and Finley, D. (2005). Delivery of ubiquitinated substrates to protein-unfolding machines. *Nat. Cell Biol.* 7, 742-749.

Feng, W.C., Southwood, C.M., and Bieker, J.J. (1994). Analyses of  $\beta$ -thalassemia mutant DNA interactions with erythroid Krueppel-like factor (EKLF), an erythroid cell-specific transcription factor. *J. Biol. Chem.* 269, 1493-1500.

Ferreon, J.C., Martinez-Yamout, M.A., Dyson, H.J., and Wright, P.E. (2009). Structural basis for subversion of cellular control mechanisms by the adenoviral E1A oncoprotein. *Proc. Natl. Acad. Sci. USA* 106, 13260-13265.

Geng, F., Wenzel, S., and Tansey, W.P. (2012). Ubiquitin and proteasomes in transcription. *Ann. Rev. Biochem.* 81, 177-201.

Grzesiek, S., and Bax, A. (1992). Correlating backbone amide and side chain resonances in larger proteins by multiple relayed triple resonance NMR. *J. Am. Chem. Soc.* 114, 6291-6293.

Haglund, K., and Dikic, I. (2005). Ubiquitylation and cell signaling. *EMBO J.* 24, 3353-3359.

Haupt, Y., Maya, R., Kazaz, A., and Oren, M. (1997). Mdm2 promotes the rapid degradation of p53. *Nature* 387, 296-299.

Hershko, A., Ciechanover, A., and Varshavsky, A. (2000). The ubiquitin system. *Nat. Med.* 6, 1073-1081.

Hochstrasser, M. (2000). Evolution and function of ubiquitin-like protein-conjugation systems. *Nat. Cell Biol.* 2, E153-157.

Hoeller, D., Crosetto, N., Blagoev, B., Raiborg, C., Tikkanen, R., Wagner, S., Kowanetz, K., Breitling, R., Mann, M., Stenmark, H., et al. (2006). Regulation of ubiquitin-binding proteins by monoubiquitination. *Nat. Cell Biol.* 8, 163-169.

Hoeller, D., Hecker, C.M., Wagner, S., Rogov, V., Dotsch, V., and Dikic, I. (2007). E3-independent monoubiquitination of ubiquitin-binding proteins. *Molecular cell* 26, 891-898.

Hurley, J.H., Lee, S., and Prag, G. (2006). Ubiquitin-binding domains. *Biochem. J.* 399, 361-372.

Ikeda, F., and Dikic, I. (2008). Atypical ubiquitin chains: new molecular signals. 'Protein Modifications: Beyond the Usual Suspects' review series. *EMBO Rep.* 9, 536-542.

Ikura, M., Kay, L.E., and Bax, A. (1991). Improved three-dimensional  $^1\text{H}$ - $^{13}\text{C}$ - $^1\text{H}$  correlation spectroscopy of a  $^{13}\text{C}$ -labeled protein using constant-time evolution. *J. Biomol. NMR* 1, 299-304.

Kay, L.E., Xu, G.Y., and Yamazaki, T. (1994). Enhanced-sensitivity triple-resonance spectroscopy with minimal  $\text{H}_2\text{O}$  saturation. *J. Magn. Reson. A*, 129-133.

Kodadek, T. (2010). No Splicing, No Dicing: Non-proteolytic Roles of the Ubiquitin-Proteasome System in Transcription. *J. Biol. Chem.* 285, 2221-2226.

Kodadek, T., Sikder, D., and Nalley, K. (2006). Keeping transcriptional activators under control. *Cell* 127, 261-264.

Koradi, R., Billeter, M., and Wuthrich, K. (1996). MOLMOL: a program for display and analysis of macromolecular structures. *J. Mol. Graphics* 14, 51-55, 29-32.

Kubbutat, M.H.G., Jones, S.N., and Vousden, K.H. (1997). Regulation of p53 stability by Mdm2. *Nature* 387, 299-303.

Langlois, C., Mas, C., Di Lello, P., Jenkins, L.M., Legault, P., and Omichinski, J.G. (2008). NMR structure of the complex between the Tfb1 subunit of TFIIH and the activation domain of VP16: structural similarities between VP16 and p53. *J. Am. Chem. Soc.* 130, 10596-10604.

Laskowski, R.A., Antoon, J., Rullmann, C., Macarthur, M.W., Kaptein, R., and Thornton, J.M. (1996). AQUA and PROCHECK-NMR: Programs for checking the quality of protein structures solved by NMR. *J. Biomol. NMR* 8, 477-486.

Le Cam, L., Linares, L.K., Paul, C., Julien, E., Lacroix, M., Hatchi, E., Triboulet, R., Bossis, G., Shmueli, A., Rodriguez, M.S., et al. (2006). E4F1 is an atypical ubiquitin ligase that modulates p53 effector functions independently of degradation. *Cell* 127, 775-788.

Lee, S., Tsai, Y.C., Mattera, R., Smith, W.J., Kostelansky, M.S., Weissman, A.M., Bonifacino, J.S., and Hurley, J.H. (2006). Structural basis for ubiquitin recognition and autoubiquitination by Rabex-5. *Nat. Struct. Mol. Biol.* 13, 264-271.

Logan, T.M., Olejniczak, E.T., Xu, R.X., and Fesik, S.W. (1992). Side chain and backbone assignments in isotopically labeled proteins from two heteronuclear triple resonance experiments. *FEBS Lett.* 314, 413-418.

Mattera, R., and Bonifacino, J.S. (2008). Ubiquitin binding and conjugation regulate the recruitment of Rabex-5 to early endosomes. *EMBO J.* 27, 2484-2494.

Mesplede, T., Gagnon, D., Bergeron-Labrecque, F., Azar, I., Senechal, H., Coutlee, F., and Archambault, J. (2012). p53 degradation activity, expression, and subcellular localization of E6 proteins from 29 human papillomavirus genotypes. *J. Virol.* 86, 94-107.

Miller, I.J., and Bieker, J.J. (1993). A novel, erythroid cell-specific murine transcription factor that binds the CACCC element and is related to the Kruppel family of nuclear proteins. *Mol. Cell. Biol.* 13, 2776-2786.

Miller, J., and Gordon, C. (2005). The regulation of proteasome degradation by multi-ubiquitin chain binding proteins. *FEBS Lett.* 579, 3224-3230.

Molinari, E., Gilman, M., and Natesan, S. (1999). Proteasome-mediated degradation of transcriptional activators correlates with activation domain potency in vivo. *EMBO J.* 18, 6439-6447.

Muratani, M., Kung, C., Shokat, K.M., and Tansey, W.P. (2005). The F box protein Dsg1/Mdm30 is a transcriptional coactivator that stimulates Gal4 turnover and cotranscriptional mRNA processing. *Cell* 120, 887-899.

Penengo, L., Mapelli, M., Murachelli, A.G., Confalonieri, S., Magri, L., Musacchio, A., Di Fiore, P.P., Polo, S., and Schneider, T.R. (2006). Crystal structure of the ubiquitin binding domains of rabex-5 reveals two modes of interaction with ubiquitin. *Cell* 124, 1183-1195.

Perkins, A.C., Sharpe, A.H., and Orkin, S.H. (1995). Lethal beta-thalassemia in mice lacking the erythroid CACCC-transcription factor EKLF. *Nature* 375, 318-322.

Pickart, C.M. (2001). Mechanisms underlying ubiquitination. *Ann. Rev. Biochem.* 70, 503-533.

Punga, T., Bengoechea-Alonso, M.T., and Ericsson, J. (2006). Phosphorylation and ubiquitination of the transcription factor sterol regulatory element-binding protein-1 in response to DNA binding. *J. Biol. Chem.* 281, 25278-25286.

Quadrini, K.J., and Bieker, J.J. (2006). EKLF/KLF1 is ubiquitinated in vivo and its stability is regulated by activation domain sequences through the 26S proteasome. *FEBS Lett.* 580, 2285-2293.

Salghetti, S.E., Caudy, A.A., Chenoweth, J.G., and Tansey, W.P. (2001). Regulation of transcriptional activation domain function by ubiquitin. *Science* 293, 1651-1653.

Salghetti, S.E., Muratani, M., Wijnen, H., Futcher, B., and Tansey, W.P. (2000). Functional overlap of sequences that activate transcription and signal ubiquitin-mediated proteolysis. *Proc. Natl. Acad. Sci. USA* 97, 3118-3123.

Sass, J., Cordier, F., Hoffmann, A., Rogowski, M., Cousin, A., Omichinski, J.G., Löwen, H., and Grzesiek, S. (1999). Purple Membrane Induced Alignment of Biological Macromolecules in the Magnetic Field. *J. Am. Chem. Soc.* 121, 2047-2055.

Shih, S.C., Katzmann, D.J., Schnell, J.D., Sutanto, M., Emr, S.D., and Hicke, L. (2002). Epsins and Vps27p/Hrs contain ubiquitin-binding domains that function in receptor endocytosis. *Nat. Cell Biol.* 4, 389-393.

Sundqvist, A., Bengoechea-Alonso, M.T., Ye, X., Lukiyanchuk, V., Jin, J., Harper, J.W., and Ericsson, J. (2005). Control of lipid metabolism by phosphorylation-dependent

degradation of the SREBP family of transcription factors by SCF(Fbw7). *Cell Metabol.* 1, 379-391.

Sundqvist, A., and Ericsson, J. (2003). Transcription-dependent degradation controls the stability of the SREBP family of transcription factors. *Proc. Natl. Acad. Sci. USA* 100, 13833-13838.

Swanson, K.A., Kang, R.S., Stamenova, S.D., Hicke, L., and Radhakrishnan, I. (2003). Solution structure of Vps27 UIM-ubiquitin complex important for endosomal sorting and receptor downregulation. *EMBO J.* 22, 4597-4606.

Ulrich, H.D. (2012). Ubiquitin and SUMO in DNA repair at a glance. *J Cell Sci* 125, 249-254.

Vijaykumar, S., Bugg, C.E., Wilkinson, K.D., and Cook, W.J. (1985). 3-Dimensional structure of ubiquitin at 2.8 Å resolution. *Proc. Natl. Acad. Sci. USA* 82, 3582-3585.

Vranken, W.F., Boucher, W., Stevens, T.J., Fogh, R.H., Pajon, A., Llinas, M., Ulrich, E.L., Markley, J.L., Ionides, J., and Laue, E.D. (2005). The CCPN data model for NMR spectroscopy: development of a software pipeline. *Proteins* 59, 687-696.

Weissman, A.M. (2001). Themes and variations on ubiquitylation. *Nat. Rev. Mol. Cell Biol.* 2, 169-178.

Winget, J.M., and Mayor, T. (2010). The diversity of ubiquitin recognition: hot spots and varied specificity. *Mol. Cell* 38, 627-635.

Wittekind, M., and Mueller, L. (1993). HNCACB, a high-sensitivity 3D NMR experiment to correlate amide-proton and nitrogen resonances with the alpha- and beta-carbon resonances in proteins. *J. Magn. Reson. B* 101, 201-205.

Wojciak, J.M., Martinez-Yamout, M.A., Dyson, H.J., and Wright, P.E. (2009). Structural basis for recruitment of CBP/p300 coactivators by STAT1 and STAT2 transactivation domains. *EMBO J.* 28, 948-958.

Yien, Y.Y., and Bieker, J.J. (2012). Functional interactions between erythroid Kruppel-like factor (EKLF/KLF1) and protein phosphatase PPM1B/PP2Cbeta. *J. Biol. Chem.* 287, 15193-15204.

Zwahlen, C., Legault, P., Vincent, S.J.F., Greenblatt, J., Konrat, R., and Kay, L.E. (1997). Methods for measurement of intermolecular NOEs by multinuclear NMR spectroscopy: application to a bacteriophage I N-peptide/boxB RNA. *J. Am. Chem. Soc.* 119, 6711-6721.



## ***SUPPLEMENTARY EXPERIMENTAL PROCEDURES***

### **Determination of K<sub>d</sub> from NMR titrations**

The apparent dissociation constants (K<sub>d</sub>) for the EKLFTAD1-UBI and EKLFTAD2-UBI complexes were determined using NMR titrations. A series of 2D <sup>1</sup>H-<sup>15</sup>N HSQC NMR experiments were performed and chemical shift changes were quantified using a weighted chemical shift change per residue Δ, where  $\Delta = [\Delta\delta_{\text{H}}^2 + (0.15 \Delta\delta_{\text{N}})^2]^{1/2}$ . For the calculations, ΔδH and ΔδN represent the changes in chemical shifts of <sup>1</sup>HN and <sup>15</sup>N (ppm), respectively, which occur during the titration. The dissociation constant for the 1:1 binding equilibrium, K<sub>d</sub>, can be estimated by fitting the observed chemical shift changes to the following equation (Fielding, 2003):

$$\Delta = \Delta_{\text{max}} (X + 1 + (K_d/[P]) - \{(X + 1 + (K_d/[P]))^2 - 4X\}^{1/2})/2$$

In which Δ is the observed chemical shift change at a given total ligand/protein ratio, X, [P] is the protein concentration and Δ<sub>max</sub> is the change in Δ observed at saturation.

## ***SUPPLEMENTAL REFERENCE***

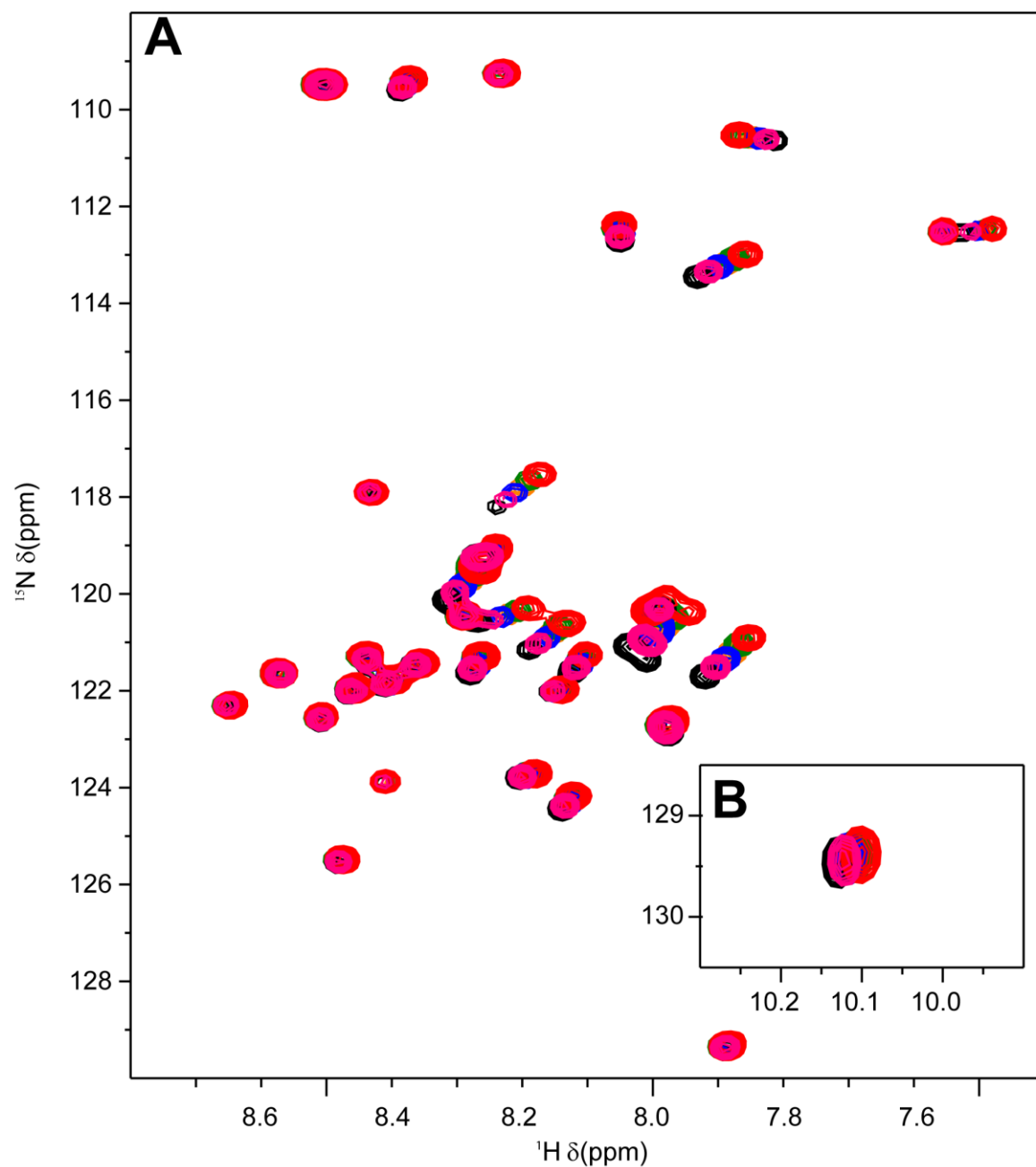
Fielding, L. (2003) NMR Methods for the Determination of Protein-Ligand Dissociation Constants. *Curr. Top. Med. Chem.* 3, 39-53.

**SUPPLEMENTARY TABLE****Tableau I. NMR and refinement statistics for EKLFTAD1-UBI complex<sup>a</sup>**

|  | <b>EKLFTAD1-UBI</b> |
|--|---------------------|
| <b>NMR distance and dihedral constraints</b>       |                     |
| Distance constraints                               |                     |
| Total NOE  | 1877                |
| Intra-residue                                      | 764                 |
| Inter-residue                                      |                     |
| Sequential ( $ i - j  = 1$ )                       | 460                 |
| Medium-range ( $ i - j  < 4$ )                     | 222                 |
| Long-range ( $ i - j  > 5$ )                       | 412                 |
| Intermolecular                                     | 19                  |
| Hydrogen bonds                                     | 38                  |
| Total dihedral angle restraints                    | 100                 |
| $\phi$   | 50                  |
| $\psi$   | 50                  |
| <b>Structure statistics</b>                        |                     |
| Violations (mean and s.d.)                         |                     |
| Distance constraints (Å)                           | 0.0106 ± 0.0006     |
| Dihedral angle constraints (°)                     | 0.226 ± 0.018       |
| Max. dihedral angle violation (°)                  | 2.2                 |
| Max. distance constraint violation (Å)             | 0.22                |
| Deviations from idealized geometry                 |                     |
| Bond lengths (Å)                                   | 0.00184 ± 0.00007   |
| Bond angles (°)                                    | 0.362 ± 0.005       |
| Impropers (°)                                      | 0.228 ± 0.010       |
| Average pairwise r.m.s. deviation (Å) <sup>b</sup> |                     |
| Heavy  | 1.30 ± 0.10         |
| Backbone   | 0.56 ± 0.10         |
| Ramachandran statistics (%) <sup>c</sup>           |                     |
| Residues in most favored regions                   | 72.2                |
| Residues in additional allowed regions             | 25.2                |
| Residues in generously allowed regions             | 2.0                 |
| Residues in disallowed regions                     | 0.6                 |

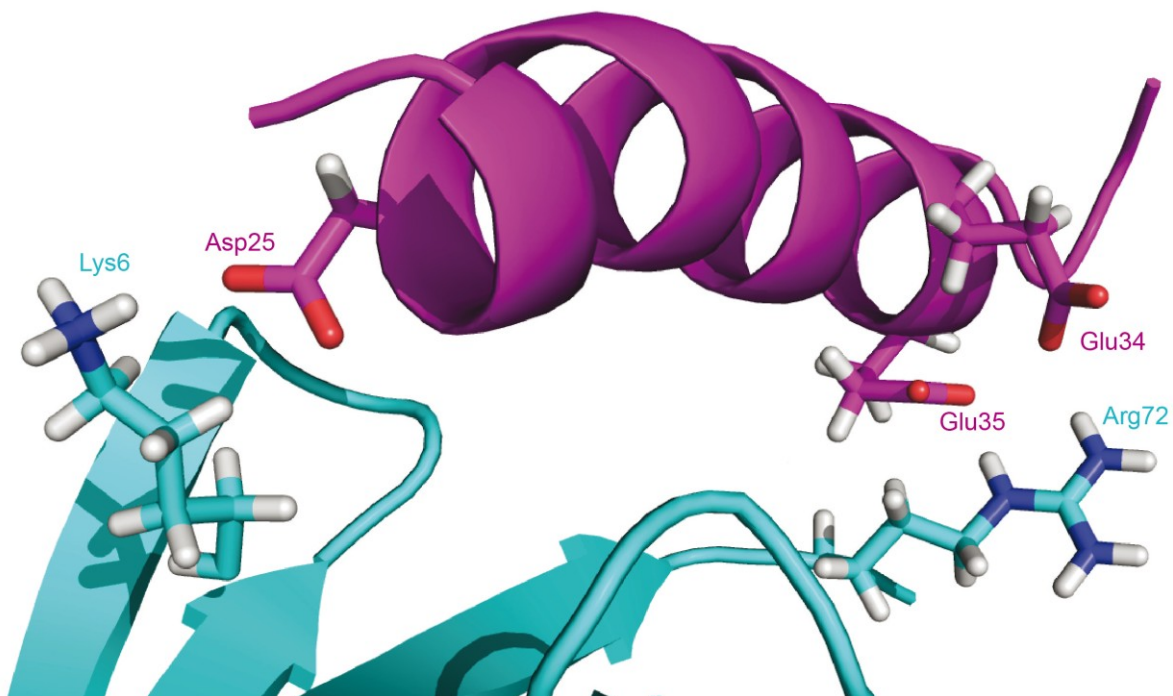
<sup>a</sup> The 20 conformers with the lowest energies were selected for statistical analysis.<sup>b</sup> Only residues 3–72 of UBI and residues 24–35 of EKLFTAD1 were used for the r.m.s.d. calculations. Residues at the N-terminus (1–3) and at the C-terminus (73–76) of UBI, as well as residues at the N-terminus (1–23) and at the C-terminus (36–40) of EKLFTAD1 were not included in the calculation.<sup>c</sup> Based on PROCHECK-NMR analysis.

**SUPPLEMENTARY FIGURES**



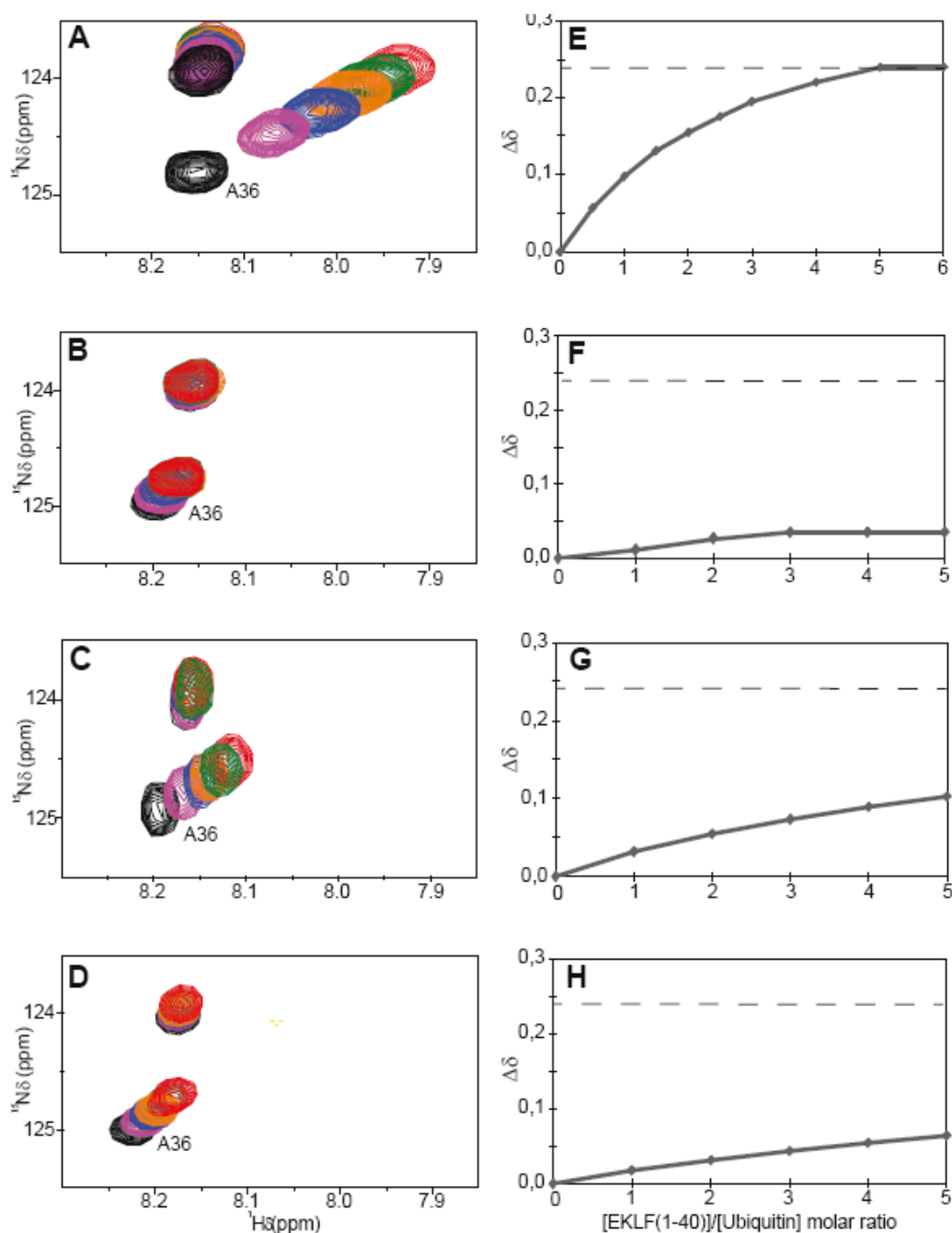
**Supplementary Figure 2.S 1 related to Figure 2. 1. EKLFTAD2 binds weakly to UBI.**

(A) Overlay of the 2D  $^1\text{H}$ - $^{15}\text{N}$  HSQC spectrum of  $^{15}\text{N}$ -labeled EKLFTAD2 in the free form (black) and in the presence of 1 (pink), 2 (blue), 3 (orange), 4 (green) and 5 (red) molar equivalents of unlabeled UBI. (B) A zoom highlighting the signal from the side chain of Trp (H $\epsilon$ 1) from the titrations with  $^{15}\text{N}$ -labeled EKLFTAD2 with UBI is shown in the small rectangle.



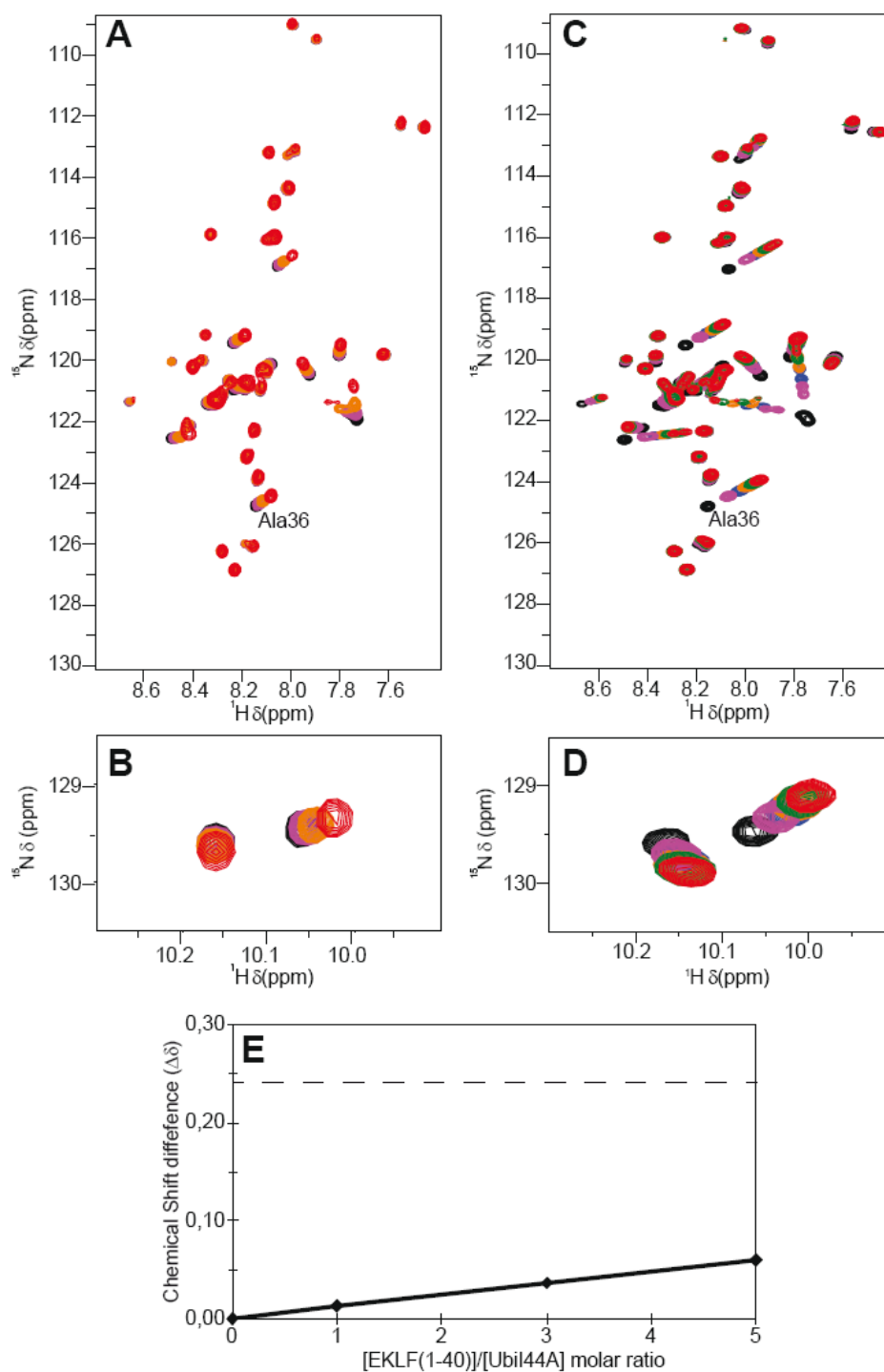
**Supplementary Figure 2.S 2 related to Figure 2. 4. Positively charged residues of UBI help to position the negatively charged EKLFTAD1.**

Ribbon representation of UBI (cyan) and EKLFTAD1 (magenta) highlighting the positively charged residues on the surface of UBI and the negatively charged residues of EKLFTAD1. In this structure, Asp25 of EKLFTAD1 is positioned near Lys6 of UBI and Glu34 and Glu35 of EKLFTAD1 are positioned near Arg72 of UBI.



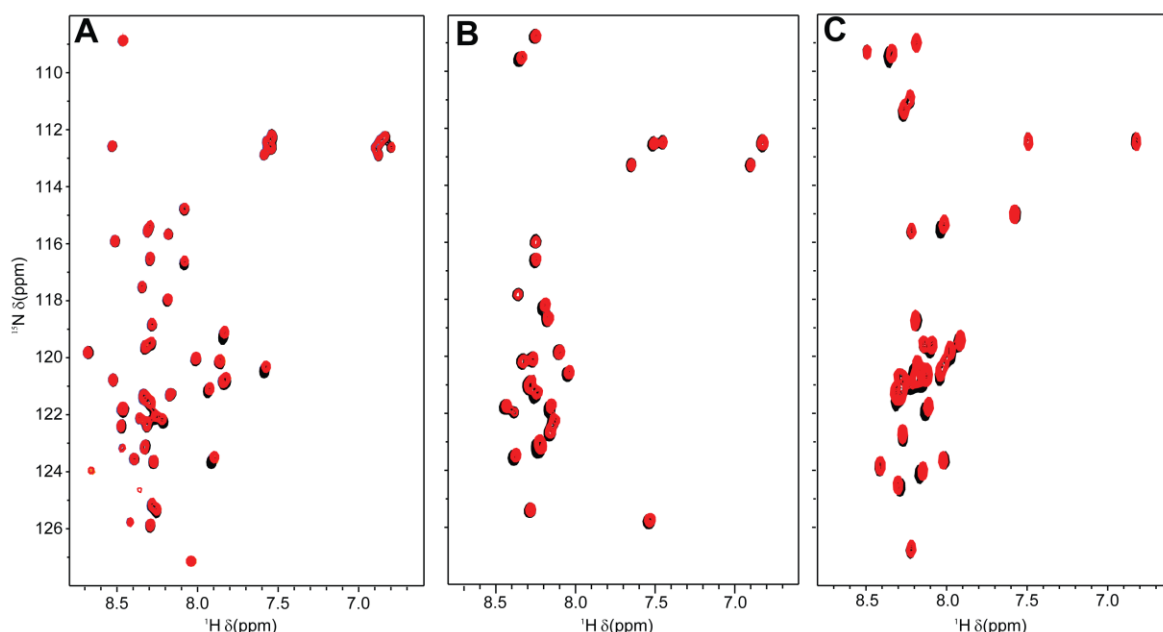
**Supplementary Figure 2.S 3 related to Figure 2. 4. Phe27, Leu28 and Trp31 from EKLFTAD1 are crucial for binding to UBI.**

Selected regions of the overlay of the 2D  $^1\text{H}$ - $^{15}\text{N}$  HSQC spectrum of  $^{15}\text{N}$ -labeled EKLFTAD1 (A) and  $^{15}\text{N}$ -labeled EKLFTAD1 variants F27S (B) L28S (C) and Trp31S (D) in the free form (black) and in the presence of 1 (pink), 2 (blue), 3 (orange), 4 (green) and 5 (red) molar equivalents of unlabeled UBI. Ala36 from either EKLFTAD1 or the EKLFTAD1 variants is highlighted in each titration to demonstrate the relative strength of binding to UBI. Weighted average chemical shift deviation ( $\Delta\delta$ ) for the amide signal of Ala37 plotted as a function of [EKLFTAD1 or EKLFTAD1 mutant]/[UBI] molar ratio for EKLFTAD1 (E), EKLFTAD1F27S (F), EKLFTAD1 L28S (G) and EKLFTAD1W31S (H). In E-H, the solid line trace fitted curves as described in the text and the dashed line is the final deviation obtained at saturation with EKLFTAD1.



**Supplementary Figure 2.S 4 related to Figure 2. 4. Ile44 of UBI is a crucial for EKLFTAD1 binding to UBI.**

Overlay of the 2D  $^1\text{H}$ - $^{15}\text{N}$  HSQC spectrum of  $^{15}\text{N}$ -labeled EKLFTAD1 in the free form (black) and in the presence of 1 (pink), 3 (orange) and 5 (red) molar equivalents of unlabeled UBI I44A mutant (A-B) or in the presence of 1 (pink), 2 (blue), 3 (orange), 4 (green) and 5 (red) molar equivalents UBI (C-D). (E) Weighted average chemical shift deviation ( $\Delta\delta$ ) for the amide signal of Ala37 plotted as a function of [EKLFTAD1]/[UBI I44A] molar ratio.



**Supplementary Figure 2.S 5 related to Figure 2. 6. p53TAD1, Gal4TAD and VP16NTAD do not bind UBI with significant affinity.**

Overlay of the 2D  $^1\text{H}$ - $^{15}\text{N}$  HSQC spectrum of  $^{15}\text{N}$ -labeled p53TAD1 (A),  $^{15}\text{N}$ -labeled Gal4TAD (B) and  $^{15}\text{N}$ -labeled VP16NTAD (C) in the free form (black) and in the presence of 5 (red) molar equivalents of unlabeled UBI. In the three titrations, only very small chemical shift changes are observed even after addition of 5 molar equivalents of UBI.

|         |           |                          |
|---------|-----------|--------------------------|
| EKLF    | (22-35)   | DTQDDFLKWWRSEE           |
| SREBP1a | (27-40)   | TDIEDMLQLINNQD           |
| P53     | (17-30)   | ETFSDLWKLLPENN           |
| Gal4    | (859-872) | TTMDDVYNYLFDDE           |
| VP16N   | (431-444) | MTPADALDDFDLEM           |
|         |           | * * $\phi$ $\phi$ $\phi$ |

**Supplementary Figure 2.S 6 related to Figure 2. 6. Sequence Homology between EKLTAD1 and other acidic TADs that function as activation-degradation regions.**

Sequence alignment of acidic TADs from transcription factors whose regions function as activation-degradation regions: EKLF, SREBP1a, p53, Gal4 and VP16N.  $\phi$  indicates a hydrophobic residue while \* indicates an acidic residue or a residue that could be potentially phosphorylated.





### **3. Discussion**



### 3.1 Retour sur les résultats

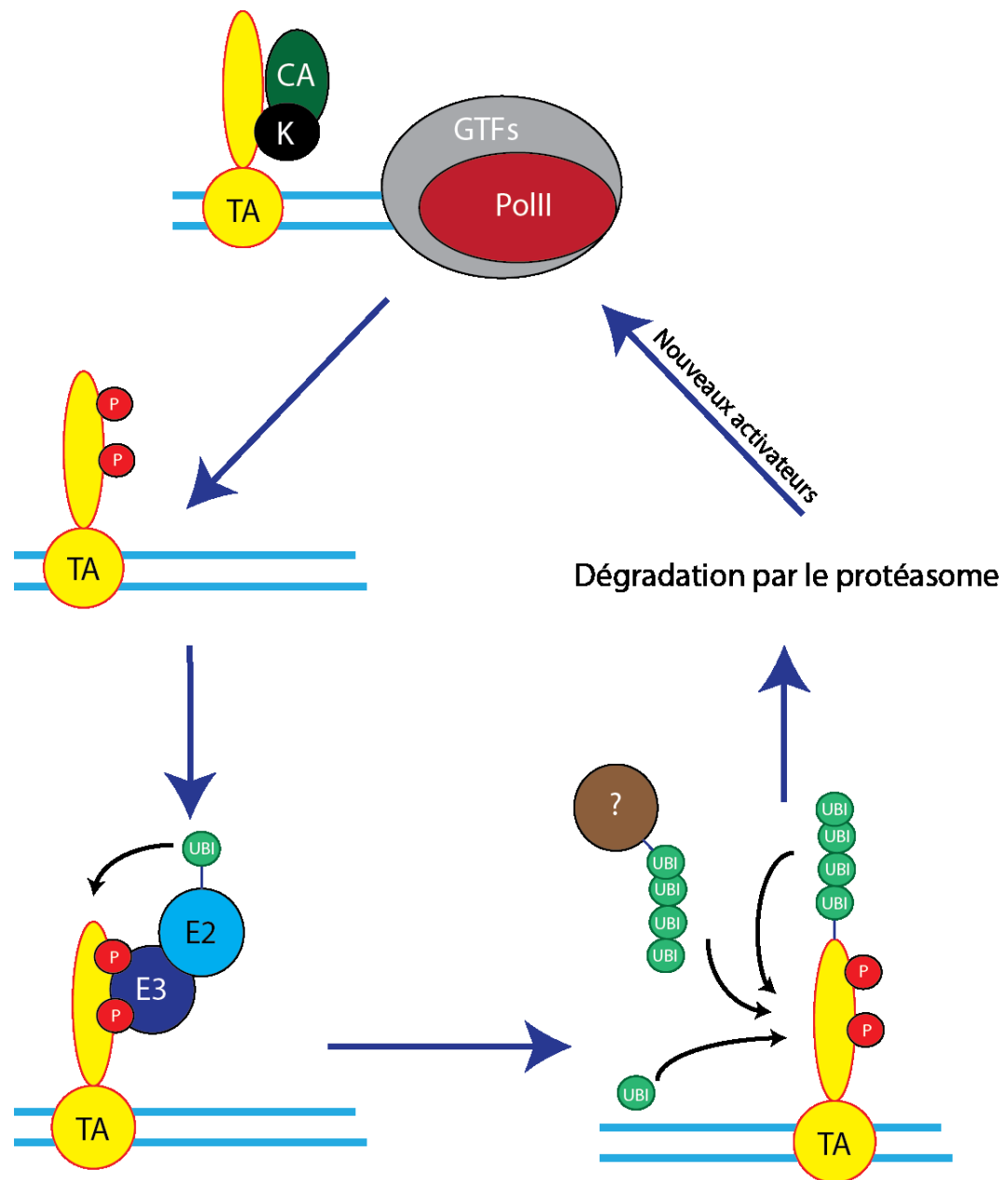
L'article présenté dans la section précédente démontre l'existence d'une interaction non-covalente entre l'UBI et le DAD du facteur de transcription EKLF (référé comme EKLFTAD1 dans l'article) comprenant les résidus 1 à 40. Il démontre aussi que les résidus responsables sur EKLF pour cette interaction avaient un impact considérable sur la dégradation de ce dernier. En effet, les mutations F27S, L28S et W31S sur EKLF empêchent la formation du complexe entre le DAD d'EKLF et l'UBI et améliorent la stabilité d'EKLF *in vivo*. D'autre part, la résolution de la structure du complexe montre un certain nombre de similitudes avec le complexe formé par l'UBI et le domaine UBD de Rabex5. Ce dernier est connu pour lier la forme ubiquitinée de récepteurs de facteurs de croissance épidermiques (EGFR : Epidermal Growth Factor Receptors) suggérant que le DAD d'EKLF pourrait partager les mêmes types de fonctions que les EGFRs.

Cet article caractérise le tout premier cas d'une interaction entre l'UBI et le DAD d'un TAD acide. Cependant, contrairement à nos attentes, cette interaction n'est pas commune pour tous les domaines DAD testés (EKLF, p53, VP16, SREBP1a, Gal4). En effet, bien que ces derniers soient tous des domaines intrinsèquement désordonnés possédant une séquence similaire, seuls les DAD d'EKLF et de SREBP1a sont capables de lier l'UBI. Ainsi, nos résultats illustrent que la capacité d'un DAD à lier l'UBI n'est pas un élément commun pour tous les sous-domaines DAD de différents AT, suggérant des mécanismes de contrôle distincts.

Les mécanismes de régulation de la fonction paradoxale à travers les DAD sont très complexes. Comme mentionné dans le chapitre 1, il a été suggéré que les DAD fonctionnent principalement selon deux modèles distincts (modèle d'activation négative et modèle d'activation positive) où l'intervention du protéasome joue un rôle crucial (**Figure 1.13** de la section 1.3.3). Le modèle de contrôle négatif des DAD, qui inclut ceux de p53 VP16 et Gal4, passe directement par leur liaison à une E3 ligase pour initier la dégradation de l'AT par le protéasome via des signaux d'ubiquitination. Celui du contrôle positif des DAD, qui inclut ceux d'EKLF et de SREBP1a, suggère l'intervention d'enzymes kinases pour initier la liaison à une E3 ligase qui initie sa dégradation par le protéasome

via l'ubiquitination. La découverte que des DAD du modèle positif soient capables d'interagir avec l'UBI et que les résidus importants pour cette interaction soient essentiels à la dégradation de l'AT va pouvoir contribuer au raffinement du modèle de régulation des DAD. Cependant, on ignore encore à quel moment de la régulation l'interaction UBI/DAD se met en place et aussi comment elle affecte la fonction des DAD.

Concernant le moment de l'interaction de l'UBI avec les DAD d'activateurs dans le modèle d'activation positive, il a été démontré que la délétion des domaines DAD chez EKLF et SREBP1a n'affecte pas leur niveau d'ubiquitination malgré la suspension de la dégradation de l'AT engendrée [116]. Ainsi, si l'on se fie au modèle de contrôle positif, ceci suggère que la formation du complexe UBI/DAD survient de façon séquentielle entre le temps de la dégradation par le protéasome et celui de l'ubiquitination par une E3 ligase (Figure 3.1). Cependant, il est aussi possible que le modèle positif soit erroné et que l'ubiquitination par la E3 ne soit pas requise pour la dégradation. En effet, de plus en plus d'études démontrent la possibilité d'une dégradation par le protéasome indépendante de signaux d'ubiquitination, et ce, surtout chez les protéines possédant des domaines intrinsèquement désordonnés comme il a été démontré pour les facteurs de transcription c-Fos et Fra-1 [128, 129]. Ainsi, l'interaction UBI/DAD pourrait être un élément qui initie ce type de dégradation. Or, ce scénario est très peu probable puisque nous avons des données expérimentales qui démontrent que la délétion du domaine responsable pour initier l'ubiquitination par la E3 ligase chez SREBP1a et EKLF est nécessaire pour leur dégradation (données non montrées). Ainsi, ceci suggère que l'intervention de l'interaction UBI/DAD et de l'ubiquitination par une E3 chez les AT, du moins pour EKLF et SREBP1a, sont des événements intimement reliés pour assurer la régulation de leur dégradation. Malgré ce lien étroit, on ignore encore comment cette interaction UBI/DAD pourrait affecter la régulation de la fonction paradoxale (**figure 3.1**).



**Figure 3. 1. Modèle proposé de la régulation des domaines paradoxaux d'EKLf et de SREBP1a.**

Le modèle schématique suggère une étape supplémentaire au modèle de régulation positive où la formation d'une interaction non-covalente entre le DAD de l'AT et l'UBI survient entre l'ubiquitination de l'AT par le recrutement d'une E3 et la dégradation par le protéasome. Dans cette nouvelle étape, l'UBI dans l'interaction non-covalente pourrait venir de l'UBI libre, de l'AT ubiquitiné lui-même ou encore de n'importe quelle autre protéine ubiquitinée. Abréviations : CA, coactivateur; E2, enzyme de conjugaison de l'UBI; E3, enzyme de ligase UBI; TA, activateur de transcription; UB, ubiquitine; P, phosphorylation; K, kinase; GTF, facteurs généraux de transcription; lignes bleues, l'ADN; ?, n'importe quelle protéine ubiquitinée.

Le but premier des domaines de liaison à l'UBI est de reconnaître des protéines modifiées par l'UBI. Ainsi, la reconnaissance de l'UBI sur les DAD pourrait servir à recruter des protéines ubiquitinées connues pour jouer un rôle dans l'activation de la transcription et/ou la dans la dégradation de protéines. La découverte de l'interaction de l'UBI avec les DAD des facteurs de transcription EKLF et de SREBP1a ouvre la porte à différentes interprétations possibles pour expliquer leur rôle au sein du modèle d'activation positive. En effet, la capacité d'interaction avec l'UBI suggère que les DAD de ce modèle peuvent interagir avec n'importe quelle protéine qui se fait ubiquitiner. Sachant que presque toutes les protéines humaines se font ubiquitiner, trouver la ou les protéines qui interviennent dans le processus de contrôle de la fonction paradoxale à travers les DAD est un projet complexe. Cependant, si on ne considère que les protéines qui peuvent affecter à la fois la dégradation et l'activation de la transcription, le nombre de candidats se restreint.

L'interaction non-covalente avec UBI peut être l'un des éléments nécessaires pour lier des éléments de remodelage de la chromatine ubiquitinés comme les histones. Ce type d'interaction permettrait aux DAD de servir comme rassembleur d'éléments essentiels pour directement réguler à la fois l'activation et la répression de gènes. En effet, plusieurs histones sont sujets à une régulation de leur activité par l'ubiquitination non-dégradative [130]. Les histones H2A et H2B, qui sont utilisées pour dérouler la chromatine et rendre les gènes accessibles pour la transcription, sont sujets à une mono-ubiquitination qui leur confère des fonctions opposées. La mono-ubiquitination de H2A est associée à la répression des gènes tandis que celle de H2B est associée à la stimulation de l'activation de la transcription. Ainsi, les UBD présents sur les DAD pourraient interagir avec des histones mono-ubiquitinées pour directement contrôler la transcription.

Il est aussi possible que l'interaction UBI/DAD serve à l'ubiquitination de l'activateur. En effet, il a été montré que des protéines contenant des UBD de type UIM sont capables d'initier le processus d'ubiquitination indépendamment d'une E3 ligase via une interaction avec une enzyme de conjugaison E2 chargée d'UBI [131]. Cependant, étant donné qu'il a été montré que la délétion du domaine DAD n'affecte pas

l'ubiquitination globale des AT du deuxième modèle testé, cette hypothèse est très peu probable [29].

Un lien plus direct avec le deuxième modèle serait que l'interaction UBI/DAD pourrait servir d'échafaud pour amener des éléments du système protéosomal qui sont impliqués autant dans la dégradation que dans l'activation de la transcription. En effet, plusieurs éléments ubiquitinés du protéasome sont présents au sein du noyau et pourraient potentiellement être dirigés vers les AT liés à l'ADN via l'interaction DAD/UBI pour initier la transcription et la dégradation [112].

L'interaction entre UBI et les DAD de EKLF et SREBP1a peut aussi être l'un des éléments nécessaires pour lier des branchements UBI spécifiques. En effet, bien que nous ayons démontré que l'interaction soit possible avec la mono-UBI, la majorité des domaines UBD ont une préférence pour un type d'embranchement poly-UBI spécifique (K48, K63, K8, etc.) [77]. Il serait intéressant de démontrer le type d'embranchement poly-UBI spécifique au DAD d'EKLF et de SREBP1a. Cette information pourrait écarter certaines des hypothèses énumérées précédemment et aider à mieux comprendre le rôle de l'affinité des DAD pour l'UBI.

La faible affinité de l'interaction UBI/DAD peut être représentative d'une interaction avec une molécule dérivée de l'UBI. En effet, les modificateurs UBI-like, bien qu'ils possèdent une séquence différente de l'UBI, possèdent une structure quasiment identique à cette dernière. Les protéines UBI-like comme « small ubiquitin-like modifier » (SUMO), « Neural Precursor Cell Expressed Developmentally Down-Regulated Protein 8 » (NEDD8) ou encore « FUnction of Boundary 1 » (FUB1) sont tous des protéines qui peuvent avoir un impact sur le contrôle de la transcription [132]. Les UBI-like peuvent aussi faire partie intégrante d'une protéine entière, par exemple celui du facteur d'élongation de la transcription B (elongin B) [133, 134]. Ce dernier possède un domaine UBI-like en N-terminal qui lui permet de jouer un rôle important dans la régulation de l'activation de la transcription. De plus, Elongin B est aussi connecté à des éléments du protéasome [112]. Ainsi, sachant que le protéasome joue un rôle essentiel dans le modèle

d'activation positive, Elongin B est un candidat intéressant qui pourrait lier les domaines DAD.

Les nombreuses hypothèses concernant le rôle précis de la région DAD de TAD acides montrent que le mécanisme est complexe et requiert la prise en compte de plusieurs éléments qui ne sont pas nécessairement communs pour tous les AT. La découverte que certains sous-domaines DAD sont capables de former un complexe avec l'UBI n'est qu'un début dans la compréhension de la fonction paradoxale des DAD. Des études additionnelles devront être effectuées pour déterminer clairement le rôle joué par l'interaction non-covalente des DAD d'EKLf et de SREBP1a avec l'UBI et de comprendre pourquoi les DAD du modèle d'activation négative n'ont pas besoin de cette interaction pour bien fonctionner.

### **3.2. Projets futurs**

#### ***3.2.1. La liaison à l'UBI et le modèle d'activation négative***

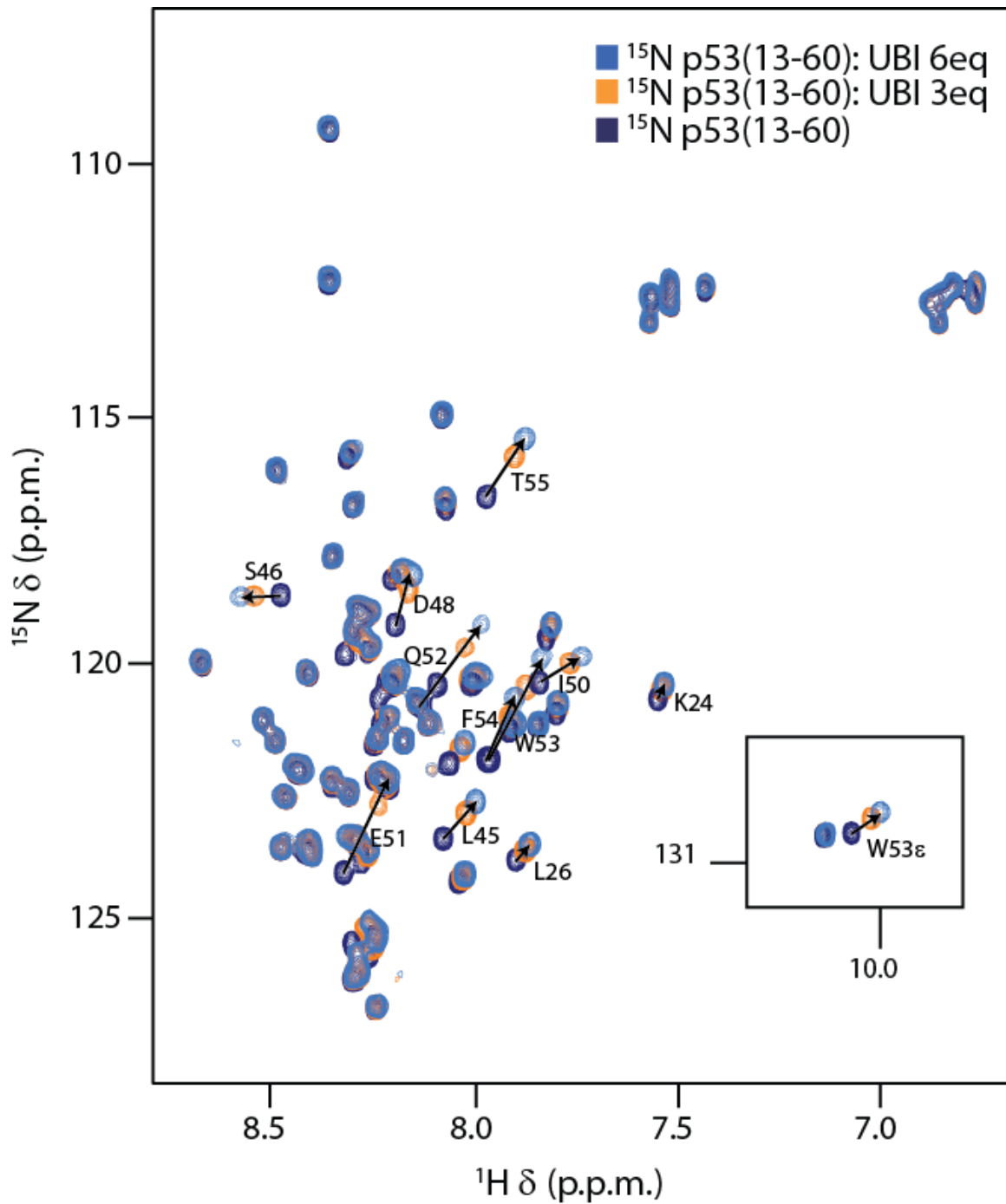
Tout au long de ce travail, nous avons porté beaucoup d'attention sur la capacité des sites DAD suivant le modèle d'activation positive à lier l'UBI. Cependant, malgré le fait que nous ayons démontré que ceux suivant le modèle d'activation négative n'étaient pas capables de lier l'UBI, cela n'indique pas nécessairement que l'interaction non-covalente avec l'UBI au niveau du TAD complet de ces AT n'existe pas. Plusieurs hypothèses suggèrent que l'interaction serait possible, du moins pour le TAD de p53. La première étant qu'il est possible que les domaines DAD de p53 dans le chapitre 2 aient été sous-estimés. Par exemple, la région DAD testée de p53 (résidus 1 à 40) ne comporte que la moitié de sa séquence riche en résidus acides et hydrophobes de son TAD complet (résidus 1 à 73), qui sont reconnus pour interagir avec l'UBI. La deuxième hypothèse se base sur le fait que p53 possèdent deux TAD distincts. Ces deux TAD peuvent avoir des fonctions redondantes sans pour autant avoir le même mécanisme. Par exemple, EKLf possède lui aussi deux TAD qui sont connus pour affecter la dégradation, cependant, seul son TAD1 (qui contient on DAD) est capable d'interagir avec l'UBI. De la même façon, les deux TAD de p53 ont un effet sur sa dégradation [135]. Cependant, les mécanismes de



cette fonction pour chacun de ces TAD sont différents puisque seul le TAD1 contrôle la stabilité de la protéine en se liant à la E3 ligase MDM2 tandis qu'aucune preuve démontre qu'une E3 ligase lie le TAD2 de p53 [136]. De ce fait, il est possible que p53 TAD2 soit relié à la dégradation par un autre moyen comme une interaction non-covalente avec l'UBI.

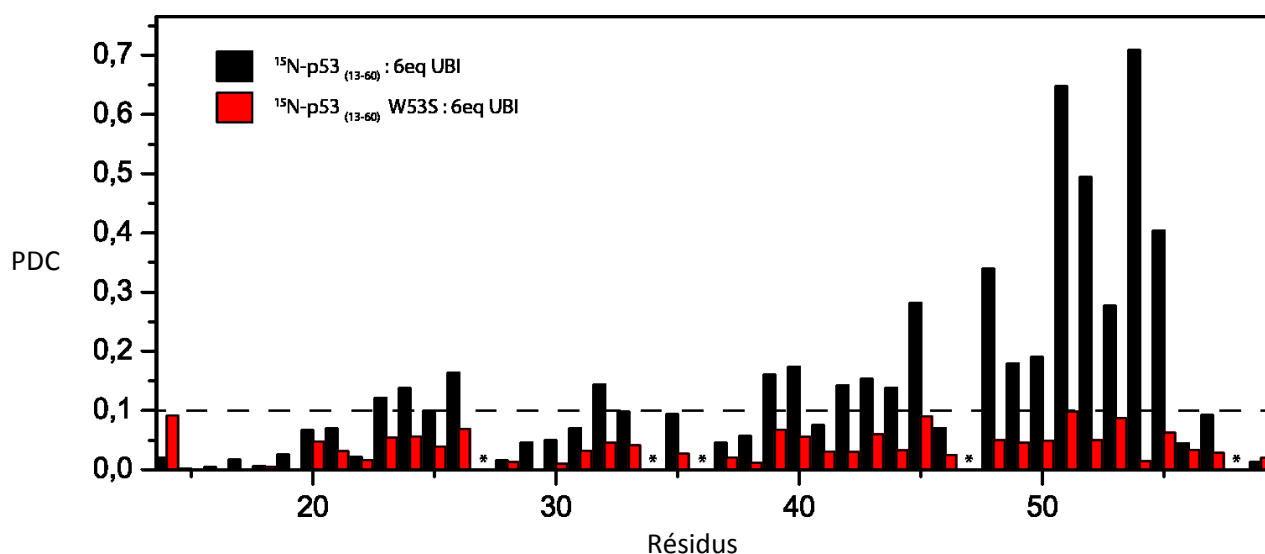
Des résultats préliminaires obtenus à partir d'expériences de titrage HSQC avec une construction du TAD complet de p53 (p53<sub>(13-60)</sub>) qui contient les résidus 13 à 60) et UBI ont permis d'identifier la présence d'un domaine UBD concentré principalement sur les résidus 48 à 55 de p53. L'addition d'UBI non-marquée à <sup>15</sup>N-p53TAD cause des changements significatifs dans la dimension <sup>1</sup>H et <sup>15</sup>N pour plusieurs signaux dans le spectre 2D <sup>1</sup>H-<sup>15</sup>N-HSQC (**Figure 3.2**). Cela suggère que malgré son incapacité à lier le DAD de p53TAD1, l'UBI peut former une interaction non-covalente qui affecte principalement des résidus sur p53TAD2. Ainsi, sachant qu'une interaction UBI/TAD peut affecter la stabilité d'un AT, il est possible que l'interaction p53<sub>(13-60)</sub>/UBI ait un impact sur sa dégradation. Des essais de dégradation avec des mutants ciblés pour défaire la formation de ce complexe devront être effectués pour enquêter sur cette possibilité.

Une étude, portée sur p53, a démontré que son résidu W53 est essentiel au bon fonctionnement de son TAD2 [135]. Ainsi, afin de déterminer si cette mutation avait un effet sur la formation du complexe p53<sub>(13-60)</sub>/UBI, nous avons effectué des expériences de titrage RMN avec p53<sub>(13-60)</sub>\_W53S et UBI. Ces expériences démontrent que cette mutation empêche la formation du complexe (**Figure 3.3**). Cependant, malgré ce résultat, il est difficile de prouver que la formation du complexe p53<sub>(13-60)</sub>/UBI soit directement lié à tous les effets engendrés par la mutation W53S puisque cette mutation est aussi connue pour empêcher l'interaction de p53 avec plusieurs facteurs impliqués dans la régulation de l'activité transcriptionnelle incluant les co-activateurs CBP/p300 et la sous-unité p62 du facteur général de transcription FTIIH [8, 10, 120, 137]. La résolution de la structure par RMN du complexe p53<sub>(13-60)</sub>/UBI pourrait nous permettre de localiser les particularités de l'interaction et de les utiliser pour faire des essais de mutations plus ciblés. Toutefois, il reste aussi à déterminer si le fait que le TAD de p53 soit capable de lier l'UBI soit une exception par rapport aux autres activateurs du modèle négatif testés dans la section précédente.



**Figure 3. 2. Le TAD de p53 et sa capacité à lier l'UBI.**

Représentation des spectres 2D  $^1\text{H}$ - $^{15}\text{N}$  HSQC de  $^{15}\text{N}$ -p53<sub>(13-60)</sub> en présence de 0 (mauve) 3 (orange) et 6 (bleu) équivalents molaires d'UBI. L'attribution des résidus de  $^{15}\text{N}$ -p53<sub>(13-60)</sub> les plus affectés par l'interaction est indiquée par les flèches.

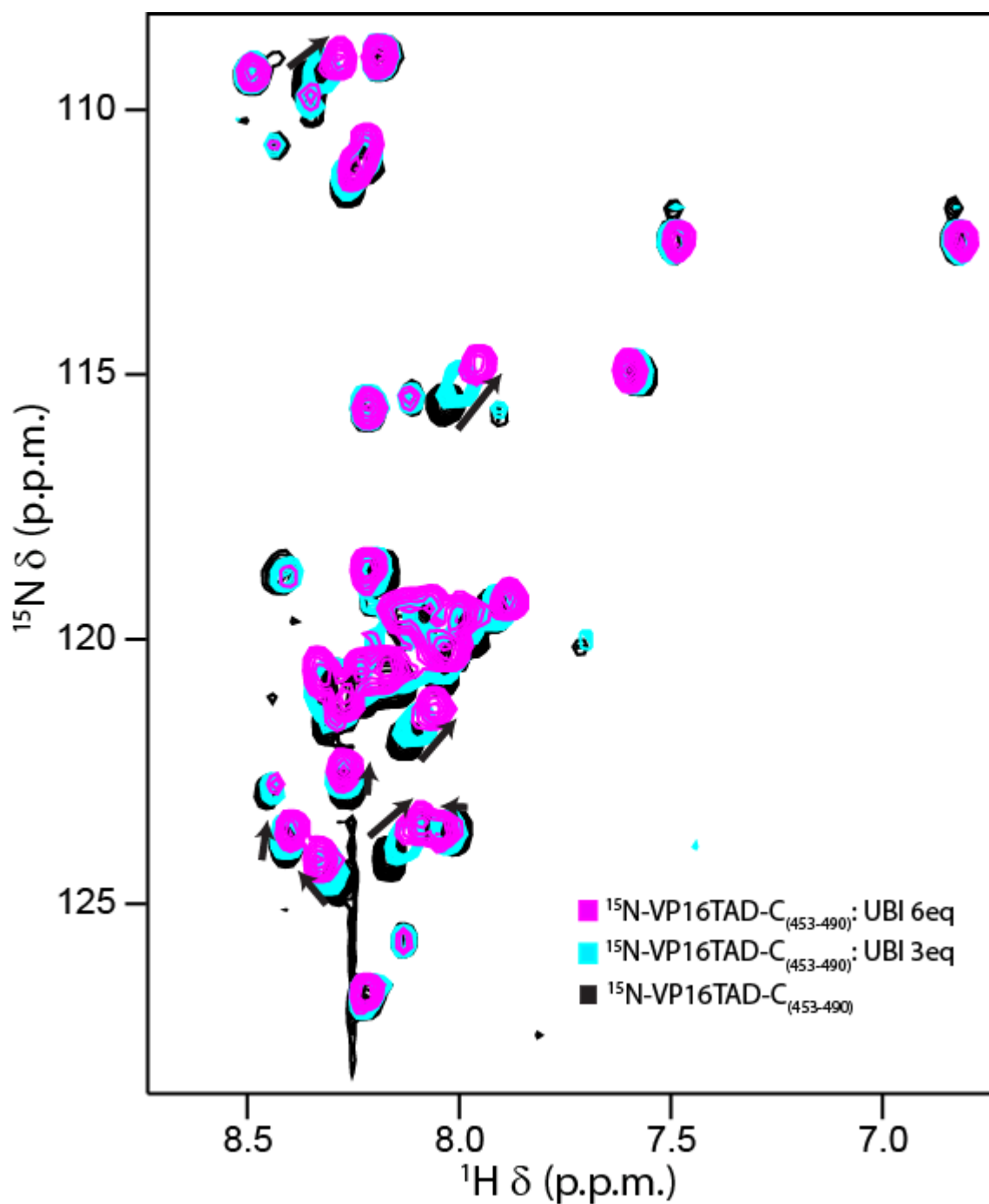


**Figure 3. 3. Effet de la mutation p53\_W53S sur la formation du complexe avec l'UBI.**

Perturbation des déplacements chimiques sur les groupements amides de  $^{15}\text{N-p53TAD}(13-60)$  et  $^{15}\text{N-p53TAD}(13-60)\text{_W53S}$  suite aux points finaux du titrage d'UBI non marqué dans le spectre 2D  $^1\text{H-}^{15}\text{N}$  HSQC. Dans la figure, les \* représentent des résidus proline et la barre pointillée représente la limite standard de perturbations chimiques considérées significatives.

### 3.2.2. La liaison à l'UBI et les TAD viraux

Les protéines virales ont évolué pour copier certaines fonctions protéiques humaines et pour en délaissier d'autres afin de pouvoir assurer leur survie au sein de l'organisme hôte. Il serait intéressant de déterminer si les TAD acides de protéines virales sont capables d'interagir avec l'UBI. Dans ce mémoire nous avons déterminé que le DAD de la protéine virale VP16 n'était pas capable de lier l'UBI. Cependant, comme p53, cet AT possède deux domaines TAD : un en N-terminal (VP16TAD-N<sub>(410-452)</sub>) qui chevauche le domaine DAD (résidus 410-452) et un en C-terminal (VP16TAD-C<sub>(453-490)</sub>) qui contient les résidus 453 à 490) [138]. Sachant que des AT possédant deux TAD, comme p53, sont probablement capables d'interagir avec l'UBI à travers leur TAD complet, il serait intéressant de vérifier si cette interaction est possible pour un TAD viral. Des expériences de titrage RMN entre <sup>15</sup>N-VP16TADC et UBI non-marquée démontrent que l'addition d'UBI non-marquée à un échantillon contenant <sup>15</sup>N-VP16TAD-C<sub>(453-490)</sub> initie des petites variations dans les dimensions <sup>1</sup>H et <sup>15</sup>N pour quelques signaux dans le spectre 2D <sup>1</sup>H-<sup>15</sup>N-HSQC (**Figure 3.4**). Ces résultats indiquent une ressemblance importante entre les TAD d'une protéine humaine (p53) et ceux d'une protéine virale (VP16). Il reste à savoir si l'interaction non-covalente entre VP16TADC<sub>(453-490)</sub> et UBI est essentielle à sa virulence. De plus, pour élargir cette conclusion à l'ensemble des protéines virales, d'autres candidats qui contiennent des TAD acides comme Varicella-zoster Virus IE62 [139], pourraient être testés.



**Figure 3. 4. Le TADC de VP16 et sa capacité à lier l'UBI.**

Représentation des spectres 2D  $^1\text{H}$ - $^{15}\text{N}$  HSQC de  $^{15}\text{N}$ -VP16TAD-C<sub>(453-490)</sub> en présence de 0 (noir) 3 (cyan) et 6 (fuchsia) équivalents molaires d'UBI. Les résidus de  $^{15}\text{N}$ -VP16TAD-C<sub>(453-490)</sub> les plus affectés par l'interaction sont indiqués par les flèches.

### 3.2.3. La liaison à l'UBI, la E3 ligase et les TAD de p53 et d'EKL

L'échéance temporelle des événements régulant un mécanisme est un élément capital pour la régulation de toutes les fonctions de la cellule. Dans les modèles du contrôle de la fonction paradoxale, la liaison de la E3 ligase semble être un élément essentiel et commun aux deux modèles [112]. Il serait intéressant de déterminer si la formation du complexe TAD/UBI vient s'incorporer avant, après ou de manière concomitante avec l'interaction d'une E3 ligase dans les deux modèles de régulation.

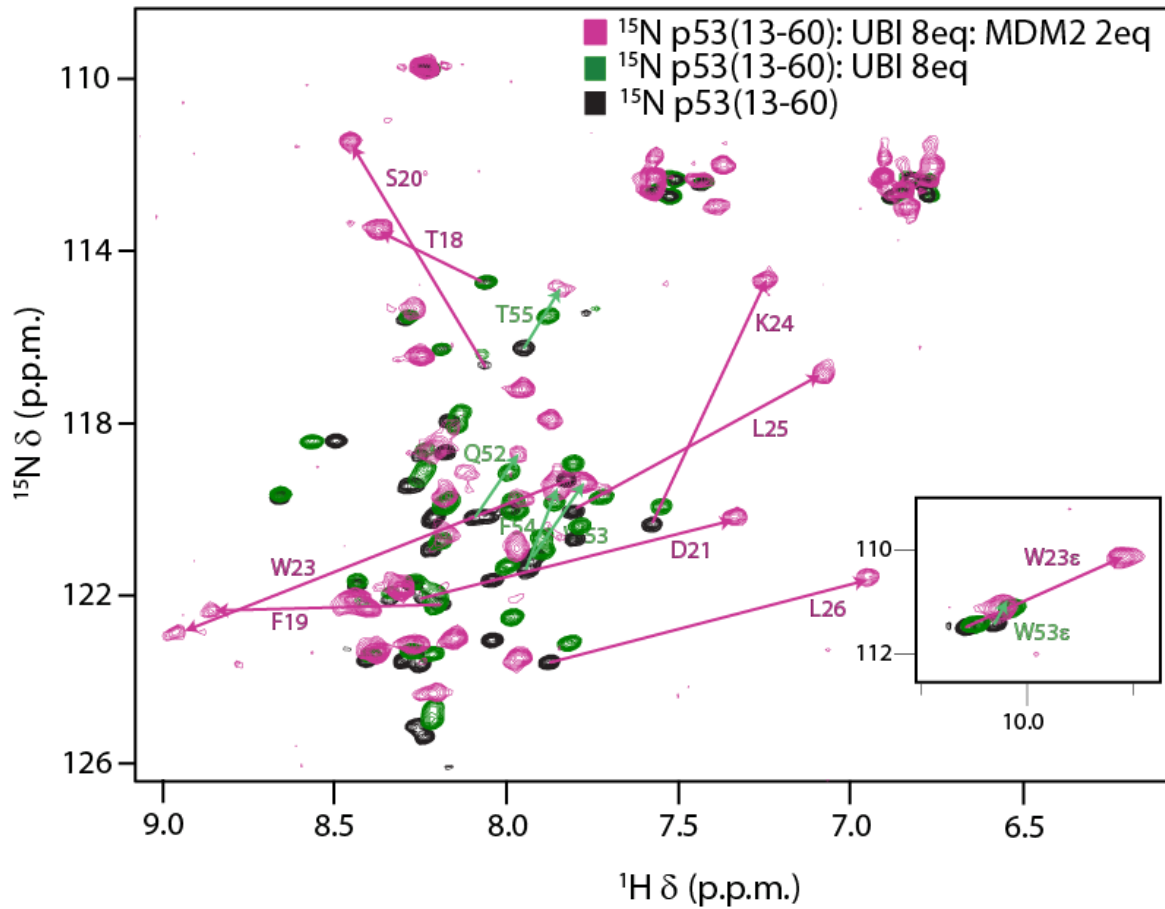
#### 3.2.3.1. p53, MDM2 et UBI

Il est intéressant de noter que la dualité entre le mécanisme d'activation par les TAD et celui de diriger l'ubiquitination est parfois inhérent à plusieurs ligases E3 elles-mêmes, où dans certains cas la E3 agit comme ubiquitine ligase et comme coactivateur transcriptionnel. Par exemple, la protéine E6AP originellement identifier come une E3 ligase qui signale la destruction de certaines protéines, est d'agir comme co-activateur pour la transcription de gènes codant pour des récepteurs de stéroïdes. Ainsi, étant donné le lien étroit entre les TAD de p53 et la E3 ligase MDM2, il est possible qu'il y ait un lien entre la dualité fonctionnelle des TAD et celle des E3.

Le TAD1 de p53 est reconnu pour être le centre de l'interaction avec la E3 ligase MDM2 qui régule la dégradation de p53. Cependant, nous savons que la mutation de W53 sur le TAD2 engendre une dégradation plus prononcée de p53 par cette même E3 ligase [135]. Étant donné que MDM2 n'est pas au contact de ce résidu lors de l'interaction, ceci indique qu'il y aurait un facteur impliqué avec le TAD2 de p53 qui influence l'efficacité de dégradation de p53 par MDM2. Ainsi, vu la relation étroite entre MDM2 et l'UBI dans le système de l'UBI, et le fait que ces deux protéines lient p53 principalement à deux endroits distincts, il est possible que ces deux éléments coopèrent pour réguler la dégradation. Des expériences préliminaires démontrent que malgré sa forte affinité avec le TAD1 de p53 (0.4  $\mu$ M) [140], le domaine de MDM2 connu pour interagir avec p53 (MDM2<sub>(17-125)</sub> qui contient les résidus 17 à 125) est incapable de déplacer la liaison de UBI au TAD2 (**Figure 3.5**). L'addition de MDM2<sub>(17-125)</sub> à <sup>15</sup>N-p53TAD cause de grands déplacements chimiques au niveau de la résonance des atomes N $\epsilon_1$ -H $\epsilon_1$  du W23 (TAD1), mais très peu sur ceux du Trp53. Pareillement, l'addition de l'UBI à ce complexe

MDM2/<sup>15</sup>N-p53TAD crée des perturbations uniquement au niveau du Nε<sub>1</sub>-Hε<sub>1</sub> de W53 (TAD2), renforçant la possibilité d'une liaison simultanée *in vitro* de MDM2 et UBI au TAD1 et TAD2, respectivement, de p53 *in vitro*. Des expériences supplémentaires de titration par RMN devront être effectuées sur des constructions marquées de MDM2 et aussi d'UBI pour confirmer la formation d'un complexe ternaire. Par ailleurs, des expériences de bioluminescence par transfert d'énergie (BRET) pourront déterminer en utilisant des fluorophores distincts sur MDM2, EKLF et sur une construction de l'UBI qui empêche sa liaison covalente à un substrat si les composants du complexe ternaire viennent à proximité l'un de l'autre dans un contexte cellulaire.

Des mutants ciblés pour défaire l'interaction entre les différents partenaires pourront ensuite être utilisés dans des essais de dégradations, comme ceux utilisés dans le chapitre 2. C'est même mutants pourront aussi être utiles dans des essais d'activation sur différents promoteurs en détectant l'activité relative d'un gène rapporteur . Ces techniques pourront déterminer si la formation du complexe non-covalent entre MDM2, UBI et les TAD de p53 a un effet sur la fonction paradoxale de ce dernier et du même coup, si leur effet est différent dépendamment du promoteur qu'il active.



**Figure 3. 5. Complexe ternaire entre p53, UBI et MDM2.**

Représentation des spectres 2D  $^1\text{H}$ - $^{15}\text{N}$  HSQC de  $^{15}\text{N}$ -p53<sub>(13-60)</sub> en présence de 0 (noir), 8 équivalents d'UBI non-marquée (vert) et dans un complexe ternaire de ratio molaire 1 : 8 : 2 avec UBI et MDM2<sub>(17-125)</sub> (rose) respectivement. Les flèches représentent la direction de la perturbation chimique de signaux correspondant aux résidus de  $^{15}\text{N}$ -p53<sub>(13-60)</sub> qui interagissent avec MDM2<sub>(17-125)</sub> (flèches roses) et avec UBI (flèches vertes).



### 3.2.3.2. EKLF, UBI et une E3 inconnue

Il n'est pas connu si EKLF est capable d'interagir avec une E3 ligase pour contrôler son activité. Cependant, le fait de savoir que cet AT se fait ubiquitiner même en présence d'inhibiteurs de protéasome laisse croire qu'il existe une E3 ligase qui assiste à l'attachement de signaux UBI pour signaler la protéolyse [29]. De plus, l'existence de cette E3 ligase est d'autant plus supportée par le fait qu'EKLF partage de nombreuses similitudes avec les mécanismes de fonction de son compatriote SREBP1a du modèle positif. SREBP1a requiert l'aide de la E3 ligase Fbw7 dans le contrôle de sa dégradation. Suite à sa liaison à l'ADN par son DBD, la section C-terminale de SREBP1a se fait phosphoryler ce qui sert ensuite de site d'ancrage afin de recruter une E3 ligase pour réguler sa stabilité [141, 142]. Il n'est pas connu si EKLF fonctionne de la même manière, cependant, nous avons effectué des expériences récentes qui suggèrent que, comme SREBP1a, son domaine DBD est essentiel à la fonction activation/dégradation du facteur de transcription (résultats non montrés). Aussi, malgré qu'il n'y ait pas de séquence de phosphorylation comme SREBP1a en C-terminal, EKLF est capable d'interagir avec une phosphatase tout près de son domaine TAD1 qui lie l'UBI [143]. Sachant que les E3 ligases ont une certaine capacité à se lier à des domaines intrinsèquement désordonnés ou mal repliés à l'aide de résidus hydrophobes exposés par ces-derniers [56, 136, 144-146], les régions TAD d'EKLF pourraient être de bons candidats comme site de liaison à une E3 ligase en plus d'être un site de liaison à l'UBI.

Il serait donc intéressant d'étudier la possibilité d'une liaison E3 ligase avec le segment désordonné complet d'EKLF qui s'étend des résidus 1 à 292. Les ressemblances entre EKLF et p53, tant au niveau de leur séquence qu'au niveau de leurs partenaires d'interaction, laissent penser qu'EKLF soit aussi capable de lier la même E3 ligase que p53 : MDM2. Si l'interaction est confirmée, un autre point d'intérêt serait de déterminer si cette liaison a un effet sur celle avec l'UBI sur le TAD1 et vérifier si ça existe *in vivo* avec les mêmes expériences suggérées dans la section 3.2.3.1. Par contre, il ne faudrait pas écarter la possibilité que d'autres E3 ligases puissent se lier sur la séquence désordonnée d'EKLF.

### **3.3. Conclusion générale**

Les TAD acides d'AT sont des domaines qui permettent la régulation de l'expression génique. Leur structure intrinsèquement désordonnée leur offre des avantages fonctionnels tels que la possibilité d'interagir avec des partenaires multiples ce qui permet notamment une meilleure régulation de leur activité. Cependant, pour ces mêmes raisons, il est difficile de prédire comment une séquence si versatile peut réguler la transcription avec une telle efficacité. L'impact de nos travaux s'avère être capital du point de vue de la compréhension des mécanismes de fonctionnement des TAD acides. En effet, nos études ont mis en évidence la diversité des mécanismes de reconnaissance des TAD cellulaires et viraux et illustrent le rôle important de l'interaction non-covalente de l'UBI dans la régulation de la stabilité des TAD de EKLF. Cependant, malgré la découverte de cette interaction avec l'UBI et de son effet, on ignore encore ce qu'elle représente au niveau de la signalisation moléculaire. Il est fort probable que cette interaction soit représentative d'une interaction avec un partenaire ubiquitiné. Par contre, nos résultats ne permettent pas de déterminer quel ou quelles sont ces partenaires. D'autres études devront être effectuées pour élucider l'identité de ces partenaires et aussi pour déterminer comment l'interaction UBI peut influencer la régulation de la stabilité des TAD de EKLF et aussi d'autres TAD capables de lier l'UBI de manière non-covalente. Néanmoins, les travaux réalisés au cours de ce mémoire apportent une contribution notable à la compréhension des mécanismes d'action des domaines TAD non structurés.

# Bibliographie

1. Blancafort, P., D.J. Segal, and C.F. Barbas, 3rd, *Designing transcription factor architectures for drug discovery*. Mol Pharmacol, 2004. **66**(6): p. 1361-71.
2. Lee, T.I. and R.A. Young, *Transcriptional regulation and its misregulation in disease*. Cell, 2013. **152**(6): p. 1237-51.
3. Hnisz, D., et al., *Super-enhancers in the control of cell identity and disease*. Cell, 2013. **155**(4): p. 934-47.
4. Orphanides, G., T. Lagrange, and D. Reinberg, *The general transcription factors of RNA polymerase II*. Genes Dev, 1996. **10**(21): p. 2657-83.
5. Garvie, C.W. and C. Wolberger, *Recognition of specific DNA sequences*. Mol Cell, 2001. **8**(5): p. 937-46.
6. Sakaguchi, K., et al., *Damage-mediated phosphorylation of human p53 threonine 18 through a cascade mediated by a casein 1-like kinase. Effect on Mdm2 binding*. J Biol Chem, 2000. **275**(13): p. 9278-83.
7. Siatecka, M., L. Xue, and J.J. Bieker, *Sumoylation of EKLF promotes transcriptional repression and is involved in inhibition of megakaryopoiesis*. Mol Cell Biol, 2007. **27**(24): p. 8547-60.
8. Langlois, C., et al., *NMR structure of the complex between the Tfb1 subunit of TFIIH and the activation domain of VP16: structural similarities between VP16 and p53*. J Am Chem Soc, 2008. **130**(32): p. 10596-604.
9. Di Lello, P., et al., *NMR structure of the amino-terminal domain from the Tfb1 subunit of TFIIH and characterization of its phosphoinositide and VP16 binding sites*. Biochemistry, 2005. **44**(21): p. 7678-86.
10. Di Lello, P., et al., *p53 and TFIIH share a common binding site on the Tfb1/p62 subunit of TFIIH*. Proc Natl Acad Sci U S A, 2008. **105**(1): p. 106-11.
11. Omichinski, J.G., et al., *NMR structure of a specific DNA complex of Zn-containing DNA binding domain of GATA-1*. Science, 1993. **261**(5120): p. 438-46.
12. Blau, J., et al., *Three functional classes of transcriptional activation domain*. Mol Cell Biol, 1996. **16**(5): p. 2044-55.
13. Hulboy, D.L. and G. Lozano, *Structural and functional analysis of p53: the acidic activation domain has transforming capability*. Cell Growth Differ, 1994. **5**(10): p. 1023-31.
14. Shen, F., et al., *Transcriptional activation domain of the herpesvirus protein VP16 becomes conformationally constrained upon interaction with basal transcription factors*. J Biol Chem, 1996. **271**(9): p. 4827-37.
15. Trouche, D. and T. Kouzarides, *E2F1 and E1A(12S) have a homologous activation domain regulated by RB and CBP*. Proc Natl Acad Sci U S A, 1996. **93**(4): p. 1439-42.
16. Mas, C., et al., *Structural and functional characterization of an atypical activation domain in erythroid Kruppel-like factor (EKLF)*. Proc Natl Acad Sci U S A, 2011. **108**(26): p. 10484-9.
17. Cheng, Y., et al., *Rational drug design via intrinsically disordered protein*. Trends Biotechnol, 2006. **24**(10): p. 435-42.
18. Gabrilove, J., *Overview: erythropoiesis, anemia, and the impact of erythropoietin*. Semin Hematol, 2000. **37**(4 Suppl 6): p. 1-3.
19. Foran, S.E. and G. Abel, *Guide to porphyrias. A historical and clinical perspective*. Am J Clin Pathol, 2003. **119** Suppl: p. S86-93.

20. Miller, I.J. and J.J. Bieker, *A novel, erythroid cell-specific murine transcription factor that binds to the CACCC element and is related to the Kruppel family of nuclear proteins*. Mol Cell Biol, 1993. **13**(5): p. 2776-86.
21. Perkins, A.C., A.H. Sharpe, and S.H. Orkin, *Lethal beta-thalassaemia in mice lacking the erythroid CACCC-transcription factor EKLF*. Nature, 1995. **375**(6529): p. 318-22.
22. Dzierzak, E. and S. Philipsen, *Erythropoiesis: development and differentiation*. Cold Spring Harb Perspect Med, 2013. **3**(4): p. a011601.
23. Siatecka, M., et al., *Severe anemia in the Nan mutant mouse caused by sequence-selective disruption of erythroid Kruppel-like factor*. Proc Natl Acad Sci U S A, 2010. **107**(34): p. 15151-6.
24. Borg, J., et al., *Haploinsufficiency for the erythroid transcription factor KLF1 causes hereditary persistence of fetal hemoglobin*. Nat Genet, 2010. **42**(9): p. 801-5.
25. Heruth, D.P., et al., *Mutation in erythroid specific transcription factor KLF1 causes Hereditary Spherocytosis in the Nan hemolytic anemia mouse model*. Genomics, 2010. **96**(5): p. 303-7.
26. Satta, S., et al., *Compound heterozygosity for KLF1 mutations associated with remarkable increase of fetal hemoglobin and red cell protoporphyrin*. Haematologica, 2011. **96**(5): p. 767-70.
27. Chen, X. and J.J. Bieker, *Erythroid Kruppel-like factor (EKLF) contains a multifunctional transcriptional activation domain important for inter- and intramolecular interactions*. EMBO J, 1996. **15**(21): p. 5888-96.
28. Rogers, S., R. Wells, and M. Rechsteiner, *Amino acid sequences common to rapidly degraded proteins: the PEST hypothesis*. Science, 1986. **234**(4774): p. 364-8.
29. Quadrini, K.J. and J.J. Bieker, *EKLF/KLF1 is ubiquitinated in vivo and its stability is regulated by activation domain sequences through the 26S proteasome*. FEBS Lett, 2006. **580**(9): p. 2285-93.
30. Essmann, F. and K. Schulze-Osthoff, *Translational approaches targeting the p53 pathway for anti-cancer therapy*. Br J Pharmacol, 2012. **165**(2): p. 328-44.
31. Levine, A.J., *p53, the cellular gatekeeper for growth and division*. Cell, 1997. **88**(3): p. 323-31.
32. Vogelstein, B., D. Lane, and A.J. Levine, *Surfing the p53 network*. Nature, 2000. **408**(6810): p. 307-10.
33. Burns, T.F. and W.S. El-Deiry, *The p53 pathway and apoptosis*. J Cell Physiol, 1999. **181**(2): p. 231-9.
34. Vousden, K.H., *Apoptosis. p53 and PUMA: a deadly duo*. Science, 2005. **309**(5741): p. 1685-6.
35. Bae, B.I., et al., *p53 mediates cellular dysfunction and behavioral abnormalities in Huntington's disease*. Neuron, 2005. **47**(1): p. 29-41.
36. Bretau, S., et al., *p53-dependent neuronal cell death in a DJ-1-deficient zebrafish model of Parkinson's disease*. J Neurochem, 2007. **100**(6): p. 1626-35.
37. Culmsee, C. and M.P. Mattson, *p53 in neuronal apoptosis*. Biochem Biophys Res Commun, 2005. **331**(3): p. 761-77.
38. Duan, W., et al., *p53 inhibitors preserve dopamine neurons and motor function in experimental parkinsonism*. Ann Neurol, 2002. **52**(5): p. 597-606.
39. Joerger, A.C. and A.R. Fersht, *Structural biology of the tumor suppressor p53 and cancer-associated mutants*. Adv Cancer Res, 2007. **97**: p. 1-23.
40. Vousden, K.H. and C. Prives, *Blinded by the Light: The Growing Complexity of p53*. Cell, 2009. **137**(3): p. 413-31.

41. Beckerman, R. and C. Prives, *Transcriptional regulation by p53*. Cold Spring Harb Perspect Biol, 2010. **2**(8): p. a000935.
42. Kim, H., et al., *p53 requires an intact C-terminal domain for DNA binding and transactivation*. J Mol Biol, 2012. **415**(5): p. 843-54.
43. Brooks, C.L. and W. Gu, *p53 ubiquitination: Mdm2 and beyond*. Mol Cell, 2006. **21**(3): p. 307-15.
44. Marine, J.C., et al., *Keeping p53 in check: essential and synergistic functions of Mdm2 and Mdm4*. Cell Death Differ, 2006. **13**(6): p. 927-34.
45. Wade, M., et al., *Hdmx modulates the outcome of p53 activation in human tumor cells*. J Biol Chem, 2006. **281**(44): p. 33036-44.
46. Zhang, Y. and H. Lu, *Signaling to p53: ribosomal proteins find their way*. Cancer Cell, 2009. **16**(5): p. 369-77.
47. Dijsselbloem, N., et al., *A critical role for p53 in the control of NF-kappaB-dependent gene expression in TLR4-stimulated dendritic cells exposed to Genistein*. J Immunol, 2007. **178**(8): p. 5048-57.
48. Vlatkovic, N., M.T. Boyd, and C.P. Rubbi, *Nucleolar control of p53: a cellular Achilles' heel and a target for cancer therapy*. Cell Mol Life Sci, 2013.
49. Shadfian, M., V. Lopez-Pajares, and Z.M. Yuan, *MDM2 and MDMX: Alone and together in regulation of p53*. Transl Cancer Res, 2012. **1**(2): p. 88-89.
50. Brady, C.A., *P53 transactivation-domain mutant mice reveal context-specific differences in P53's mechanism of action*. 2011, Stanford University: Stanford University.
51. Langlois, C., et al., *Structure-based design of a potent artificial transactivation domain based on p53*. J Am Chem Soc, 2012. **134**(3): p. 1715-23.
52. Collins, J., *Synthetic Biology: Bits and pieces come to life*. Nature, 2012. **483**(7387): p. S8-10.
53. Gaudreau, P., et al., *Structure-function studies of peptides inhibiting the ribonucleotide reductase activity of herpes simplex virus type I*. J Med Chem, 1992. **35**(2): p. 346-50.
54. Bernstein, F.C., et al., *The Protein Data Bank: a computer-based archival file for macromolecular structures*. J Mol Biol, 1977. **112**(3): p. 535-42.
55. Salghetti, S.E., et al., *Functional overlap of sequences that activate transcription and signal ubiquitin-mediated proteolysis*. Proc Natl Acad Sci U S A, 2000. **97**(7): p. 3118-23.
56. Salghetti, S.E., et al., *Regulation of transcriptional activation domain function by ubiquitin*. Science, 2001. **293**(5535): p. 1651-3.
57. Lipford, J.R., et al., *A putative stimulatory role for activator turnover in gene expression*. Nature, 2005. **438**(7064): p. 113-6.
58. Wang, X., et al., *Proteolytic instability and the action of nonclassical transcriptional activators*. Curr Biol, 2010. **20**(9): p. 868-71.
59. Bhat, K.P. and S.F. Greer, *Proteolytic and non-proteolytic roles of ubiquitin and the ubiquitin proteasome system in transcriptional regulation*. Biochim Biophys Acta, 2011. **1809**(2): p. 150-5.
60. Geng, F., S. Wenzel, and W.P. Tansey, *Ubiquitin and proteasomes in transcription*. Annu Rev Biochem, 2012. **81**: p. 177-201.
61. Behuliak, M., et al., *Ubiquitin--the kiss of death goes Nobel. Will you be quitting?* Bratisl Lek Listy, 2005. **106**(3): p. 93-100.
62. Hicke, L., H.L. Schubert, and C.P. Hill, *Ubiquitin-binding domains*. Nat Rev Mol Cell Biol, 2005. **6**(8): p. 610-21.
63. Hershko, A. and A. Ciechanover, *The ubiquitin system*. Annu Rev Biochem, 1998. **67**: p. 425-79.
64. Hochstrasser, M., *Biochemistry. All in the ubiquitin family*. Science, 2000. **289**(5479): p. 563-4.

65. Pickart, C.M., *Ubiquitin enters the new millennium*. Mol Cell, 2001. **8**(3): p. 499-504.
66. Komander, D., M.J. Clague, and S. Urbe, *Breaking the chains: structure and function of the deubiquitinases*. Nat Rev Mol Cell Biol, 2009. **10**(8): p. 550-63.
67. Sompallae, R., et al., *Epstein-barr virus encodes three bona fide ubiquitin-specific proteases*. J Virol, 2008. **82**(21): p. 10477-86.
68. Wang, S., et al., *Herpes simplex virus 1 ubiquitin-specific protease UL36 inhibits beta interferon production by deubiquitinating TRAF3*. J Virol, 2013. **87**(21): p. 11851-60.
69. Kumar, N.S.a.V., *Oncogenic viruses: DUBbing their way to cancer*. Virology Discovery, 2013. **1**(5).
70. Anderson, D.M. and D.W. Frank, *Five mechanisms of manipulation by bacterial effectors: a ubiquitous theme*. PLoS Pathog, 2012. **8**(8): p. e1002823.
71. Ye, Y. and M. Rape, *Building ubiquitin chains: E2 enzymes at work*. Nat Rev Mol Cell Biol, 2009. **10**(11): p. 755-64.
72. Behrends, C. and J.W. Harper, *Constructing and decoding unconventional ubiquitin chains*. Nat Struct Mol Biol, 2011. **18**(5): p. 520-8.
73. Ikeda, F. and I. Dikic, *Atypical ubiquitin chains: new molecular signals*. 'Protein Modifications: Beyond the Usual Suspects' review series. EMBO Rep, 2008. **9**(6): p. 536-42.
74. *The Protein Data Bank*. Methods Biochem Anal, 2003. **44**: p. 181-98.
75. Xu, P., et al., *Quantitative proteomics reveals the function of unconventional ubiquitin chains in proteasomal degradation*. Cell, 2009. **137**(1): p. 133-45.
76. Saeki, Y., et al., *Lysine 63-linked polyubiquitin chain may serve as a targeting signal for the 26S proteasome*. EMBO J, 2009. **28**(4): p. 359-71.
77. Dikic, I., S. Wakatsuki, and K.J. Walters, *Ubiquitin-binding domains - from structures to functions*. Nat Rev Mol Cell Biol, 2009. **10**(10): p. 659-71.
78. Young, P., et al., *Characterization of two polyubiquitin binding sites in the 26 S protease subunit 5a*. J Biol Chem, 1998. **273**(10): p. 5461-7.
79. Hofmann, K. and L. Falquet, *A ubiquitin-interacting motif conserved in components of the proteasomal and lysosomal protein degradation systems*. Trends Biochem Sci, 2001. **26**(6): p. 347-50.
80. Bilodeau, P.S., et al., *The Vps27p Hse1p complex binds ubiquitin and mediates endosomal protein sorting*. Nat Cell Biol, 2002. **4**(7): p. 534-9.
81. Polo, S., et al., *A single motif responsible for ubiquitin recognition and monoubiquitination in endocytic proteins*. Nature, 2002. **416**(6879): p. 451-5.
82. Raiborg, C., et al., *Hrs sorts ubiquitinated proteins into clathrin-coated microdomains of early endosomes*. Nat Cell Biol, 2002. **4**(5): p. 394-8.
83. Shih, S.C., et al., *Epsins and Vps27p/Hrs contain ubiquitin-binding domains that function in receptor endocytosis*. Nat Cell Biol, 2002. **4**(5): p. 389-93.
84. Hofmann, K. and P. Bucher, *The UBA domain: a sequence motif present in multiple enzyme classes of the ubiquitination pathway*. Trends Biochem Sci, 1996. **21**(5): p. 172-3.
85. Bertolaet, B.L., et al., *UBA domains of DNA damage-inducible proteins interact with ubiquitin*. Nat Struct Biol, 2001. **8**(5): p. 417-22.
86. Wilkinson, C.R., et al., *Proteins containing the UBA domain are able to bind to multi-ubiquitin chains*. Nat Cell Biol, 2001. **3**(10): p. 939-43.
87. Grabbe, C. and I. Dikic, *Functional roles of ubiquitin-like domain (ULD) and ubiquitin-binding domain (UBD) containing proteins*. Chem Rev, 2009. **109**(4): p. 1481-94.
88. Hurley, J.H., S. Lee, and G. Prag, *Ubiquitin-binding domains*. Biochem J, 2006. **399**(3): p. 361-72.
89. Winget, J.M. and T. Mayor, *The diversity of ubiquitin recognition: hot spots and varied specificity*. Mol Cell, 2010. **38**(5): p. 627-35.

90. Beal, R.E., et al., *The hydrophobic effect contributes to polyubiquitin chain recognition*. Biochemistry, 1998. **37**(9): p. 2925-34.
91. Mueller, T.D. and J. Feigon, *Solution structures of UBA domains reveal a conserved hydrophobic surface for protein-protein interactions*. J Mol Biol, 2002. **319**(5): p. 1243-55.
92. Chen, Z.J. and L.J. Sun, *Nonproteolytic functions of ubiquitin in cell signaling*. Mol Cell, 2009. **33**(3): p. 275-86.
93. Raasi, S., et al., *Binding of polyubiquitin chains to ubiquitin-associated (UBA) domains of HHR23A*. J Mol Biol, 2004. **341**(5): p. 1367-79.
94. Kang, R.S., et al., *Solution structure of a CUE-ubiquitin complex reveals a conserved mode of ubiquitin binding*. Cell, 2003. **113**(5): p. 621-30.
95. Lee, J.T. and W. Gu, *The multiple levels of regulation by p53 ubiquitination*. Cell Death Differ, 2010. **17**(1): p. 86-92.
96. Voges, D., P. Zwickl, and W. Baumeister, *The 26S proteasome: a molecular machine designed for controlled proteolysis*. Annu Rev Biochem, 1999. **68**: p. 1015-68.
97. Sauer, R.T. and T.A. Baker, *AAA+ proteases: ATP-fueled machines of protein destruction*. Annu Rev Biochem, 2011. **80**: p. 587-612.
98. Kodadek, T., *No Splicing, no dicing: non-proteolytic roles of the ubiquitin-proteasome system in transcription*. J Biol Chem, 2010. **285**(4): p. 2221-6.
99. Wilkinson, K.D., *Ubiquitination and deubiquitination: targeting of proteins for degradation by the proteasome*. Semin Cell Dev Biol, 2000. **11**(3): p. 141-8.
100. Beck, F., et al., *Near-atomic resolution structural model of the yeast 26S proteasome*. Proc Natl Acad Sci U S A, 2012. **109**(37): p. 14870-5.
101. Hoppe, T., et al., *Activation of a membrane-bound transcription factor by regulated ubiquitin/proteasome-dependent processing*. Cell, 2000. **102**(5): p. 577-86.
102. Wu, R.C., et al., *SRC-3 coactivator functional lifetime is regulated by a phospho-dependent ubiquitin time clock*. Cell, 2007. **129**(6): p. 1125-40.
103. Burgdorf, S., P. Leister, and K.H. Scheidtmann, *TSG101 interacts with apoptosis-antagonizing transcription factor and enhances androgen receptor-mediated transcription by promoting its monoubiquitination*. J Biol Chem, 2004. **279**(17): p. 17524-34.
104. Bres, V., et al., *A non-proteolytic role for ubiquitin in Tat-mediated transactivation of the HIV-1 promoter*. Nat Cell Biol, 2003. **5**(8): p. 754-61.
105. van der Horst, A., et al., *FOXO4 transcriptional activity is regulated by monoubiquitination and USP7/HAUSP*. Nat Cell Biol, 2006. **8**(10): p. 1064-73.
106. Moren, A., et al., *Differential ubiquitination defines the functional status of the tumor suppressor Smad4*. J Biol Chem, 2003. **278**(35): p. 33571-82.
107. Archer, C.T., et al., *Activation domain-dependent monoubiquitylation of Gal4 protein is essential for promoter binding in vivo*. J Biol Chem, 2008. **283**(18): p. 12614-23.
108. Wilcox, A.J. and J.D. Laney, *A ubiquitin-selective AAA-ATPase mediates transcriptional switching by remodelling a repressor-promoter DNA complex*. Nat Cell Biol, 2009. **11**(12): p. 1481-6.
109. Wang, B., H. Suzuki, and M. Kato, *Roles of mono-ubiquitinated Smad4 in the formation of Smad transcriptional complexes*. Biochem Biophys Res Commun, 2008. **376**(2): p. 288-92.
110. Laney, J.D. and M. Hochstrasser, *Ubiquitin-dependent degradation of the yeast Mat(alpha)2 repressor enables a switch in developmental state*. Genes Dev, 2003. **17**(18): p. 2259-70.
111. Ramakrishna, S., et al., *Lys-63-specific deubiquitination of SDS3 by USP17 regulates HDAC activity*. J Biol Chem, 2011. **286**(12): p. 10505-14.
112. Muratani, M. and W.P. Tansey, *How the ubiquitin-proteasome system controls transcription*. Nat Rev Mol Cell Biol, 2003. **4**(3): p. 192-201.

113. Lipford, J.R. and R.J. Deshaies, *Diverse roles for ubiquitin-dependent proteolysis in transcriptional activation*. Nat Cell Biol, 2003. **5**(10): p. 845-50.
114. Thomas, D. and M. Tyers, *Transcriptional regulation: Kamikaze activators*. Curr Biol, 2000. **10**(9): p. R341-3.
115. Geng, F. and W.P. Tansey, *Similar temporal and spatial recruitment of native 19S and 20S proteasome subunits to transcriptionally active chromatin*. Proc Natl Acad Sci U S A, 2012. **109**(16): p. 6060-5.
116. Sundqvist, A. and J. Ericsson, *Transcription-dependent degradation controls the stability of the SREBP family of transcription factors*. Proc Natl Acad Sci U S A, 2003. **100**(24): p. 13833-8.
117. Swanson, K.A., et al., *Solution structure of Vps27 UIM-ubiquitin complex important for endosomal sorting and receptor downregulation*. EMBO J, 2003. **22**(18): p. 4597-606.
118. Wysocka, J. and W. Herr, *The herpes simplex virus VP16-induced complex: the makings of a regulatory switch*. Trends Biochem Sci, 2003. **28**(6): p. 294-304.
119. Trainor, C.D., et al., *GATA-1 associates with and inhibits p53*. Blood, 2009. **114**(1): p. 165-73.
120. Di Lello, P., et al., *Structure of the Tfb1/p53 complex: Insights into the interaction between the p62/Tfb1 subunit of TFIH and the activation domain of p53*. Mol Cell, 2006. **22**(6): p. 731-40.
121. Chandrudu, S., P. Simerska, and I. Toth, *Chemical methods for peptide and protein production*. Molecules, 2013. **18**(4): p. 4373-88.
122. Papaneophytou, C.P. and G. Kontopidis, *Statistical approaches to maximize recombinant protein expression in Escherichia coli: A general review*. Protein Expr Purif, 2013. **94C**: p. 22-32.
123. Schumann, F.H., et al., *Combined chemical shift changes and amino acid specific chemical shift mapping of protein-protein interactions*. J Biomol NMR, 2007. **39**(4): p. 275-89.
124. Carlomagno, T., *Ligand-target interactions: what can we learn from NMR?* Annu Rev Biophys Biomol Struct, 2005. **34**: p. 245-66.
125. Lecoq, L., *Etude par RMN des L,D-transpeptidases bactériennes: structure, dynamique et caractérisation de leur inhibition par les B-lactames*, in *Ecole Doctorale Chimie et Science du Vivant*. 2012, Université de Grenoble: Grenoble. p. 253.
126. Fielding, L., *NMR methods for the determination of protein-ligand dissociation constants*. Curr Top Med Chem, 2003. **3**(1): p. 39-53.
127. Chabot RP, R.L., Lussier-Price M, Morse T, Arseneault G, Archambault J and Omichinski JG. , *Structural and Functional Characterization of a Complex between the Acidic Transactivation Domain of EBNA2 and the Tfb1/p62 subunit of TFIH*. Plos Pathogenes, 2014. **Accepted for publication**.
128. Basbous, J., et al., *Ubiquitin-independent- versus ubiquitin-dependent proteasomal degradation of the c-Fos and Fra-1 transcription factors: is there a unique answer?* Biochimie, 2008. **90**(2): p. 296-305.
129. Erales, J. and P. Coffino, *Ubiquitin-independent proteasomal degradation*. Biochim Biophys Acta, 2014. **1843**(1): p. 216-21.
130. Weake, V.M. and J.L. Workman, *Histone ubiquitination: triggering gene activity*. Mol Cell, 2008. **29**(6): p. 653-63.
131. Hoeller, D., et al., *E3-independent monoubiquitination of ubiquitin-binding proteins*. Mol Cell, 2007. **26**(6): p. 891-8.
132. van der Veen, A.G. and H.L. Ploegh, *Ubiquitin-like proteins*. Annu Rev Biochem, 2012. **81**: p. 323-57.



133. Conaway, J.W., et al., *Transcription factor SIII: a novel component of the RNA polymerase II elongation complex*. Cell Mol Biol Res, 1993. **39**(4): p. 323-9.
134. Aso, T., et al., *Elongin (SIII): a multisubunit regulator of elongation by RNA polymerase II*. Science, 1995. **269**(5229): p. 1439-43.
135. Venot, C., et al., *Definition of a p53 transactivation function-deficient mutant and characterization of two independent p53 transactivation subdomains*. Oncogene, 1999. **18**(14): p. 2405-10.
136. Kussie, P.H., et al., *Structure of the MDM2 oncoprotein bound to the p53 tumor suppressor transactivation domain*. Science, 1996. **274**(5289): p. 948-53.
137. Grossman, S.R., *p300/CBP/p53 interaction and regulation of the p53 response*. Eur J Biochem, 2001. **268**(10): p. 2773-8.
138. Vojnic, E., et al., *Structure and VP16 binding of the Mediator Med25 activator interaction domain*. Nat Struct Mol Biol, 2011. **18**(4): p. 404-9.
139. Yamamoto, S., et al., *Analysis of the varicella-zoster virus IE62 N-terminal acidic transactivating domain and its interaction with the human mediator complex*. J Virol, 2009. **83**(12): p. 6300-5.
140. Ferreón, J.C., et al., *Cooperative regulation of p53 by modulation of ternary complex formation with CBP/p300 and HDM2*. Proc Natl Acad Sci U S A, 2009. **106**(16): p. 6591-6.
141. Punga, T., M.T. Bengoechea-Alonso, and J. Ericsson, *Phosphorylation and ubiquitination of the transcription factor sterol regulatory element-binding protein-1 in response to DNA binding*. J Biol Chem, 2006. **281**(35): p. 25278-86.
142. Bengoechea-Alonso, M.T. and J. Ericsson, *A phosphorylation cascade controls the degradation of active SREBP1*. J Biol Chem, 2009. **284**(9): p. 5885-95.
143. Yien, Y.Y. and J.J. Bieker, *Functional interactions between erythroid Kruppel-like factor (EKLF/KLF1) and protein phosphatase PPM1B/PP2Cbeta*. J Biol Chem, 2012. **287**(19): p. 15193-204.
144. Asamitsu, K., et al., *RING finger protein AO7 supports NF-kappaB-mediated transcription by interacting with the transactivation domain of the p65 subunit*. J Biol Chem, 2003. **278**(29): p. 26879-87.
145. Narayan, V., et al., *A multiprotein binding interface in an intrinsically disordered region of the tumor suppressor protein interferon regulatory factor-1*. J Biol Chem, 2011. **286**(16): p. 14291-303.
146. Fredrickson, E.K., et al., *Exposed hydrophobicity is a key determinant of nuclear quality control degradation*. Mol Biol Cell, 2011. **22**(13): p. 2384-95.



